DISSERTATION

SPACE COMMUNICATIONS RESPONSIVE TO EVENTS ACROSS MISSIONS (SCREAM): AN INVESTIGATION OF NETWORK SOLUTIONS FOR TRANSIENT SCIENCE SPACE SYSTEMS

Submitted by

Christopher J. Roberts

Department of Systems Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2022

Doctoral Committee:

Advisor: Thomas H. Bradley

Ronald M. Sega
John M. Borky
Steven C. Reising

ABSTRACT

SPACE COMMUNICATIONS RESPONSIVE TO EVENTS ACROSS MISSIONS (SCREAM):

AN INVESTIGATION OF NETWORK SOLUTIONS FOR TRANSIENT SCIENCE SPACE SYSTEMS

The National Academies have prioritized the pursuit of new scientific discoveries using diverse and temporally coordinated measurements from multiple ground and space-based observatories. Networked communications can enable such measurements by connecting individual observatories and allowing them to operate as a cohesive and purposefully designed system. Timely data flows across terrestrial and space communications networks are required to observe transient scientific events and processes. Currently, communications to space-based observatories experience large latencies due to manual service reservation and scheduling procedures, intermittent signal coverage, and network capacity constraints. If space communications network latencies could be reduced, new discoveries about dynamic scientific processes could be realized. However, science mission and network planners lack a systematic framework for defining, quantifying and evaluating timely space data flow implementation options for transient scientific observation scenarios involving multiple ground and space-based observatories. This dissertation presents a model-based systems engineering approach to investigate and develop network solutions to meet the needs of transient science space systems.

First, a systematic investigation of the current transient science operations of the National Aeronautics and Space Administration’s (NASA) Tracking and Data Relay Satellite (TDRS) space data network and the Neil Gehrels Swift Observatory resulted in a formal architectural model for

transient science space systems. Two methods individual missions may use to achieve timely network services were defined, quantitatively modeled, and experimentally compared.

Next, the architectural model was extended to describe two alternative ways to achieve timely and autonomous space data flows to multiple space-based observatories within the context of a purposefully designed transient science observation scenario. A quantitative multipoint space data flow modeling method based in queueing theory was defined. General system suitability metrics for timeliness, throughput, and capacity were specified to support the evaluation of alternative network data flow implementations. A hypothetical design study was performed to demonstrate the multipoint data flow modeling method and to evaluate alternative data flow implementations using TDRS. The merits of a proposed future TDRS broadcast service to implement multipoint data flows were quantified and compared to expected outcomes using the as-built TDRS network.

Then, the architectural model was extended to incorporate commercial network service providers. Quantitative models for Globalstar and Iridium short messaging data services were developed based on publicly available sources. Financial cost was added to the set of system suitability metrics. The hypothetical design study was extended to compare the relative suitability of the as-built TDRS network with the commercial Globalstar and Iridium networks.

Finally, results from this research are being applied by NASA missions and network planners. In 2020, Swift implemented the first automated command pipeline, increasing its expected gravitational wave follow-up detection rate by greater than 400%. Current NASA technology initiatives informed by this research will enable future space-based observatories to become interoperable sensing devices connected by a diverse ecosystem of network service providers.

ACKNOWLEDGEMENTS

I would like to express my profound respect and sincere appreciation for the people who serve the institutions that have made my journey possible. I thank my parents Tom and Carmel Roberts for their steadfast caring and commitment, which has allowed me to pursue my dreams. I am grateful for the lifelong friendships with my siblings, Becky Cundiff and Andrew Roberts, and Paul and Mary Ann Coons. I thank Willie and Linda for welcoming me into the Higa family and enriching my life. I am grateful for the inspiring teachers, coaches, and mentors who have shared their expertise with me and encouraged my growth. I am thankful for the many guides I have encountered, people from different walks of life, who have instilled in me a reverent wonder for the immensity and complexity of our world. I could not possibly list everyone who has helped me reach this milestone. If you are reading this page, there is a good chance you had a role.

I could not have completed this dissertation without the love and devotion of my wife and partner Beth Higa Roberts. We have navigated our family through countless vicissitudes together, not least among them co-teaching virtual kindergarten during a pandemic. Thank you also for serving admirably as a sounding board for my research ideas and as chief editor of this dissertation. All remaining sentences using passive voice are mine alone.

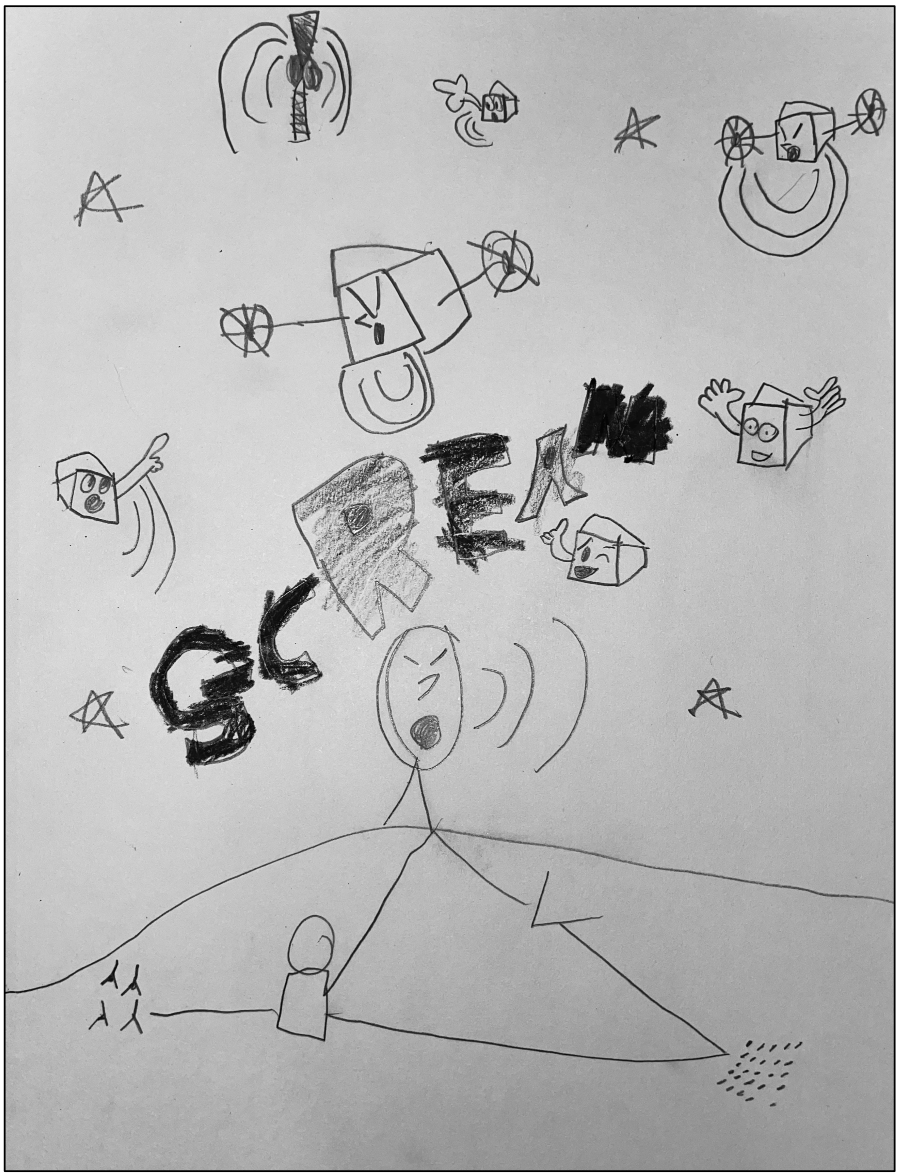
I am grateful for the dedication and unwavering support of my advisor, Prof. Tom Bradley. Thank you for your patience, guidance, and insights over the many years it took to realize the vision for this dissertation. I thank my committee members, Dr. Mike Borky, Prof. Ron Sega and Prof. Steven Reising, for contributing their expertise to my research and for their service to

Colorado State University. I thank Adam Ross, Donna Rhodes and John Thomas for supporting my application to the Colorado State systems engineering doctoral program.

I thank Dave Thomas, Mo Madden, Ken Rehm, Leslye Boyce, Gary Smith, and the Derleth Pixton, Diller, Pickworth-Burkett, Hsu and Chai families for helping me get my start and supporting my family at NASA. I am grateful for the encouragement and resources provided by leaders at the Goddard Space Flight Center: La Vida Cooper, Dave Israel, Rick Sabatino, David Matusow, David Carter, Dave Larsen, Kendall Mauldin, Ted Sobchak, Bob Menrad, Peter Hughes, Deborah Amato, Michael Johnson and Christyl Johnson. I thank Harry Shaw, Tom Gitlin, Jacob Burke, Greg Heckler, Cheryl Gramling, John Hudiburg, Rob Tye, Neil Mallik, Markland Benson, Frank Stocklin, Pat Boldosser, Robin Krause, Paul Buchanan, Jenny Donaldson, Risha George, J.P. Chamoun, Rosemary Kioko, Eric Poole, Kevin Fisher, Mark Bollard, Jake Barnes, Julie Hoover, Jimmy Acevedo, Katie Schauer, Jessica Hinkle, Kendall Murphy, Danny Baird, Dwayne Johnson, Marc Sanchez-Net, Philip Baldwin, Jay Gao, Jim Schier, Wallace Tai, Jason Mitchell, Dave Chelmins, Adam Gannon, Dale Mortensen, Janette Briones, Rob Morgenstern, Jon Verville, Wes Eddy, Juan Crenshaw, Adam Michaels, Jessica Knizhnik, Iona Rus, Jena Garrahy, Ryan Archer, Jeff Lubelczyk, Bill Horne, George Bussey, Jerry Mason, Steve Sirotzky, Mark Sinkiat, Erica Weir, Michael Hopper, Elias Alstead, Dave Richardson and the Swift & TDRS flight operations teams for their technical expertise and assistance. I thank the scientists Scott Barthelmy, Frank Marshall, Brad Cenko, Jamie Kennea, Aaron Tohuvavohu, Leo Singer, Josh Schlieder, Judy Racusin, Jeremy Perkins, Rita Sambruna and Jim Garvin for providing inspiration and facilitating practical opportunities to apply my research.

DEDICATION

This dissertation is dedicated to the memory of my grandmother Dorothy Webb and my aunt Mary Ruth Roberts in honor of their courage and perseverance. Per aspera ad astra.



SCREAM, illustrated by Nathan and Willa Higa Roberts, ages 7 and 6

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The study of transient scientific phenomena using diverse ground and space-based observatories present new opportunities for scientific breakthroughs [1] [2]. Many types of transient scientific events occur randomly and require prompt follow-up observations by multiple observatories to obtain data of the highest scientific value [3] [4] [5].

Networked communications allow multiple observatories to operate as a cohesive and purposefully designed system [6] [7] [8]. Currently, communications to space-based observatories experience large latencies due to manual service reservation and scheduling procedures, intermittent signal coverage, and network capacity constraints [9]. If space communications network latencies could be reduced, new discoveries about dynamic scientific processes could be realized [10]. However, science mission and network planners lack a systematic framework for defining, quantifying and evaluating timely space data flow implementation options for transient scientific observation scenarios involving multiple ground and space-based observatories. This dissertation presents a model-based systems engineering approach to investigate and develop network solutions to meet the needs of transient science space systems.

Dissertation Outline

Chapter 1 introduces the dissertation research and provides an outline of the document. Chapter 2 provides a literature review and presents the research questions addressed in the dissertation. Chapter 3 develops and demonstrates two complimentary space communications network service access methods to address time-sensitive mission needs using the Neil Gehrels Swift Observatory and the Tracking and Data Relay Satellite (TDRS) space data network. Chapter 4 develops two functional reference architectures to extend publish-subscribe-based terrestrial transient science data flows to space-based observatories, defines a design and evaluation method to evaluate the suitability of candidate network solutions, and demonstrates the method in a hypothetical design study of a transient science space system using the TDRS space data network. Chapter 5 extends the transient science space system architecture models, applies the design and evaluation method to the commercial Iridium and Globalstar networks and compares the relative suitability of two commercial network messaging services to TDRS for a hypothetical transient science space system design study. Chapter 6 presents the conclusions of this research.

Chapter 2

State of the Field

The current state of transient science operations and space data network solution formulation and planning practices are discussed in this section.

Current Transient Science Operations

There is growing demand for timely, collaborative, multi-observatory operations concepts across NASA's astrophysics, heliophysics, Earth science and planetary science mission domains [10]. In astrophysics, a new window on the universe has been opened with the advent of gravitational wave and high-energy neutrino detections made possible by new classes of ground-based observatories [11] [12] [13]. The astrophysical processes originating these messengers are poorly understood, but timely and diverse measurements from multiple space and ground-based observatories can lead to greater insights than would otherwise be possible [1] [14] [15].

Timely bidirectional data flows with space-based observatories are needed to realize the full potential of ground and space-based transient science observations. The TDRS Demand Access System has adequate capacity to provide continuous low-latency communications from multiple space-based observatories to the ground, so that other observatories may follow-up on space-based observations of transient events. However, there is no present means for quickly sending data from the ground so that one or more space-based observatories may follow-up on events detected by other ground or space-based observatories. Instead, to conduct follow-up operations, mission operators must wait for the next pre-scheduled network service period to

send new commands and data products to their space-based observatories or call network operators on the telephone to manually arrange a new service period.

NASA's Neil Gehrels Swift Observatory, operational since 2005, is a "first-of-its-kind multi-wavelength observatory" conceived for the study of Gamma Ray Bursts (GRB) and other transient astrophysical events and processes [4] [16] [17]. The Swift Observatory is comprised of three complimentary instruments on a single Low Earth Orbiting platform. The Burst Alert Telescope (BAT) has a field of view of approximately one sixth of the sky at a time. The BAT detects and localizes the celestial coordinates of a GRB source to within 1-3 arcminutes. Information about the GRB event source is rapidly transmitted to Swift's ground operators using the TDRS Demand Access System and then to the scientific community via a terrestrial science messaging application known as the Gamma Ray Burst Coordinates Network (GCN). The total latency, from GRB detection onboard Swift to the dissemination of data products to the scientific community via the GCN occurs within 20 seconds. Swift's flight software autonomously determines whether to engage in follow-up observations of the transient GRB source with its other onboard instruments based on the relative priority of its pre-planned activities, the relative angle between the GRB source and the sun, and other constraints. If Swift's flight software determines that a follow-up observation should be made, it will perform a slewing maneuver so that the transient source location is within the narrower field of view of its X-Ray Telescope (XRT) and Ultraviolet and Optical Telescope (UVOT) instruments. During follow-up operations, refined localization data and preliminary science data products at the XRT and UVOT instrument observing bands are disseminated via TDRS and the GCN. The stream of information about the transient source delivered by Swift allows scientists and operators for other space and ground-based

observatories determine whether or not they will perform follow-up observations according to their local priorities and constraints.

Today, the GCN has approximately 650 contributing scientists at observatories worldwide who exchanging information about transient sources using over 120 notification types. Participants may choose to subscribe to the GCN by specifying the topic, method, format and frequency of notification.

The scientific community may also submit requests to the Swift science duty officer to observe transient sources identified by their own observatories or to test hypotheses formulated based on information they have gained through the GCN. Swift receives approximately 1,400 such “target of opportunity” (TOO) requests per year. These requests are categorized according to urgency and priority, with the most urgent requests needing observation within a few minutes to less than four hours. If ground-to-space data transport latencies could be reduced, new discoveries about dynamic scientific processes could be realized.

The GCN uses a common terrestrial information technology messaging design pattern known as publish-subscribe to enable multi-observatory transient science operations [18]. The publish-subscribe design pattern allows a sender to broadcast a message to all interested receivers. It allows publishers (senders) to categorize messages into classes (GCN notification types), which can then be sent to subscribers (receivers) without the need for publishers to know the specific network routing and addressing details of the subscribers. Subscribers provide these details along with their subscription preferences to a central registry (e.g., the GCN). Some ground-based optical, radio and other observatories have implemented automated machine-to-

machine interfaces with the GCN minimize follow-up observation latency. Research is needed to explore how the publish-subscribe design pattern can be applied to space data networks.

Current Network Service Formulation and Planning Practices

The Goddard Space Flight Center's Near Space Network (NSN) provides a single point of contact for space communications and navigation network service planning and operations for space missions venturing up to two million kilometers from the Earth [19]. The NSN serves as a network service broker, matching user needs with the capabilities offered by the commercial market when possible, and with government owned assets, such as TDRS, when necessary. NSN mission users are diverse, and include NASA, other U.S. government agencies, commercial and international robotic and human mission users.

Scientists and mission planners initiate contact with NSN in a variety of ways to discuss their network service needs. During the mission formulation phases, mission concepts and their network needs are often closely held, as there are typically several mission proposal teams vying for selection. Because all missions need space communications network infrastructure, knowledge firewalls are placed between network planners working with different proposal teams. However, the resulting fragmentation of knowledge about mission needs can lead to missed opportunities for coalition building and advocacy for new capabilities. Additionally, knowledge fragmentation can lead to sub-optimal and inconsistent network solution recommendations from individual network planners.

Network feasibility assessments begin with network planners gathering information from mission stakeholders, such as the science objectives, orbit parameters and space communications data flow needs [20]. Network planners analyze the feasibility of satisfying

mission needs within the constraints of current communications resources, including spectrum availability and infrastructure capabilities and capacity. If the initial mission needs cannot be satisfied, NSN network planners propose potential trade-offs and alternatives, including augmentations to the network infrastructure, until a viable mission concept emerges. NSN maintains robust and validated space communications modeling and simulation tools to support network feasibility assessments. NSN performs approximately 90 assessments per year, at differing degrees of fidelity, as mission proposals mature over time. However, these tools have limitations for their applicability to multi-observatory transient science operations concepts.

First, the tools are primarily oriented towards predicting physical and link layer behaviors for radio frequency interfaces between a single space-based observatory and a network service provider node [21]. Transient science observation scenarios involve multipoint data flows across terrestrial and space-based networks to initiate, coordinate and execute follow-up operations. Presently unexamined factors, such network configuration wait times, are required to evaluate the feasibility of multi-observatory transient science scenarios.

Second, the network loading tools, which assess the expected increase in network utilization and cross-mission resource contention during the projected mission operational lifespan, assume mission user service needs are independent. This is valid for traditional science mission concepts, which typically involve steady-state data collection activities. However, transient science scenarios have corelated user service needs, as multiple observatories respond to a transient event.

Third, a major limitation of current feasibility assessments is that mission concepts and network architecture representations are commonly developed on an ad hoc basis using static

and informal representations in tools such as Microsoft Word, PowerPoint and Excel. Such an unstructured approach can result in disconnects between the mission concept and network service formulation process and downstream system development lifecycle activities.

In summary, there is a need to improve the NSN's current network feasibility assessment process to incorporate unique aspects of multi-observatory transient science scenarios using a more rigorous approach to system architecture.

Model-Based Systems Engineering Overview

Systems engineering involves the purposeful formulation, implementation, operations and sustainment of technical systems to address human and societal needs [22]. During formulation, systems engineers communicate recursively with stakeholders to assess needs, perform analyses and evaluate multifaceted trade-offs [23]. The design and evaluation of multi-observatory transient science scenarios and candidate network solutions spans multiple engineering disciplines, involves economic and other factors which cross managerial domains of control and requires predictions about the behaviors of systems that may not yet exist.

A mature system concept is specified by a baseline set of requirements and their traceability to both a concept of operations (comprised of a set of discrete scenarios) and an architectural representation of the system [23]. Prior to establishing this baseline, changes to system elements may occur rapidly as alternatives are explored and evaluated. Microsoft Word, PowerPoint and Excel are prevalent tools for describing system structure, behaviors, and relationships. These tools do not enforce syntactic or semantic rules or manage relationships among the instances of elements and their relationships represented in each tool. As a result,

traditional system representations have limited reusability, high variability across engineering practitioners, and ambiguous or poorly defined syntax and semantics [24].

Model-based systems engineering (MBSE) is an approach that promises several advantages over the more common ad-hoc approach to system representation [24]. MBSE leverages object-oriented methods and technologies that emerged from the discipline of software engineering. Specifically, a profile of the Unified Modeling Language (UML) known as Systems Modeling Language (SysML) has been developed to meet the needs of the systems engineering community. To carry out MBSE, a modeling language, modeling method and modeling tool are needed [25]. SysML is gaining acceptance as a more generalized systems engineering language, and there is an active commercial modeling tool market [26]. SysML provides a syntactically concise and semantically rich visual vocabulary for describing system structure, behaviors, requirements, allocations and parametric through a set of interconnected and extensible diagrams. These attributes ensure architectural traceability, consistency and rigorous specification of system representations at all levels in the system hierarchy. MBSE provides a promising means to address the limitations of current space communications network feasibility assessments.

Establishing credible multi-observatory transient science scenarios, including their relative scientific merit and the innovative potential of the enabling technologies, is a pre-requisite for competitive selection and funding for further technology and system development in the NASA systems engineering lifecycle [23] [27]. In addition to the limitations of current feasibility assessment tools discussed previously, the tools are also oriented towards computing relatively high-fidelity responses for point design solutions. The model configuration set-up and

computational run-time for a response prediction varies based on the novelty and complexity of the problem at hand but is not likely to scale well under conditions involving large numbers of observatories interacting in complicated ways. A relatively quick and simple method to predict and understand the relative merits and costs of alternative candidate network solutions is needed to explore the systems engineering tradespace for transient science space systems.

Changes to the space sector are occurring due to new and emerging commercial entrants in launch services, satellite manufacturing and operations services, and other space-based services. These market forces are expected to increase access to space, resulting in many more space platforms and, consequently, greater demand for space communications network services [28]. Prior work has established the need for improved space mission data timeliness for a broad range of important latency-sensitive applications [29] [30]. In addition, there is growing recognition for the need to consider scenarios involving distributed satellite systems, whereby functionality traditionally allocated to a single monolithic platform is instead allocated among multiple spacecraft that interact in order to achieve desired goals [6]. Multi-observatory transient science space systems present one example of an emerging class of latency-sensitive distributed satellite systems.

As a result of these trends, there is a need for a more comprehensive and systematic approach to understand and manage the complexity involved in engineering complex information-intensive aerospace systems. MBSE offers an improved means for translating broad architecture visions into tractable engineering problems in the concept development lifecycle phase. MBSE provides reusable and extensible system representations to reduce variability

among engineers, ensures architectural traceability for analysis, and provides continuity for development in downstream lifecycle phases.

Research Questions

The overarching question for this research is:

How can space data network solutions be designed and evaluated for multi-observatory transient science space systems?

To address the overarching research question, three component research questions have been defined and decomposed into tasks.

Research Question #1

What are the factors that significantly influence data latency and network traffic for a transient science space system using the TDRS space data network?

The tasks described in Table 1 are performed to systematically investigate the contributors to data latency and network traffic for a transient science space system involving a single space-based observatory using the TDRS space data network.

Table 1: Summary of tasks for Research Question #1.

Task	Description
1.1	Develop and analyze a descriptive architectural model for transient science space systems.
1.2	Develop quantitative models for data latency and network traffic.
1.3	Apply the quantitative models to a real-world transient science space system.
1.4	Analyze and interpret results.

Research Question #2

How can the publish-subscribe messaging design pattern be applied to space data networks and its impact be evaluated for transient science space systems?

The tasks described in Table 2 build on the results of Research Question #1 and address transient science space systems involving more than one space-based observatory. Alternative multipoint data flow solutions using the TDRS space data network are investigated in a hypothetical design study of a transient science space system.

Table 2: Summary of tasks for Research Question #2.

Task	Description
2.1	Extend the architectural model from Task 1.1 to incorporate publish-subscribe behaviors for space data flows.
2.2	Develop quantitative models in accordance with the architectural changes in Task 2.1.
2.3	Apply the quantitative models to a hypothetical transient science space system design study.
2.4	Analyze and interpret results.

Research Question #3

How do commercial service provider network solutions compare to TDRS for transient science space systems?

The tasks described in Table 3 build on results of Research Question #1 and #2 to incorporate commercial network service providers. Results for commercial network providers are compared to those of TDRS in an extension of the hypothetical design study from Research Question #2.

Table 3: Summary of tasks for Research Question #3.

Task	Description
3.1	Extend the architectural model from Task 2.1 to incorporate commercial network service providers.
3.2	Develop quantitative models of commercial networks based on publicly available data sources and in accordance with the architectural changes in Task 3.1.
3.3	Apply the commercial network quantitative models to a hypothetical transient science space system design study.
3.4	Analyze and interpret results.

Chapter 3

Chapter 3 addresses Research Question #1:

What are the factors that significantly influence data latency and network traffic for a transient science space system using the TDRS space data network?

This work was published in the American Institute of Aeronautics and Astronautics Journal of Aerospace Information Systems in June 2021 as "Evaluation of Timely Communications Access Methods using NASA Space Network."

Introduction

There is growing interest in developing collaborative and time-sensitive science operations concepts involving multiple distributed observatories within the National Aeronautics and Space Administration's (NASA) astrophysics, heliophysics, Earth and planetary science mission domains [6] [7]. For example, a new window on the universe has been opened in astrophysics with the advent of gravitational wave and high-energy neutrino measurements made possible by new classes of ground-based observatories [11] [14]. The scientific processes originating these observables are poorly understood, but diverse and temporally responsive measurements from multiple distributed space and ground-based observatories provide greater insights than would otherwise be possible [14] [15]. A summary of the state-of-the-practice for NASA's transient science operations is provided in this section, examining the Neil Gehrels Swift Observatory's use of network services as a motivating case study. Additionally, this section discusses the research strategy and chapter organization.

Background

Launched in 2004, the Neil Gehrels Swift Observatory is a first-of-its-kind multi-wavelength observatory created for the study of gamma-ray bursts (GRB) and other transient astrophysical sources [16]. The Swift observatory is comprised of three complimentary instruments on a single low-Earth orbiting (LEO) spacecraft [16]. Swift's innovative concept of operations is enabled by a direct-to-earth ground station network and a space relay network. High-fidelity science data is collected and stored onboard the observatory from both pre-planned observation targets and from randomly occurring transient scientific sources. Downlink of this data occurs at fixed pre-scheduled intervals to the ground station network via an X-band link service [31]. Concurrently, the ground network provides lower data rate bi-directional S-band link services between the observatory and the mission operations center (MOC) for tracking, telemetry and command data. This data is used to operate, maintain and sustain the observatory. The ground network provides intermittent connectivity to Swift due to limitations associated with ground network coverage, loading and the observatory's orbit geometry.

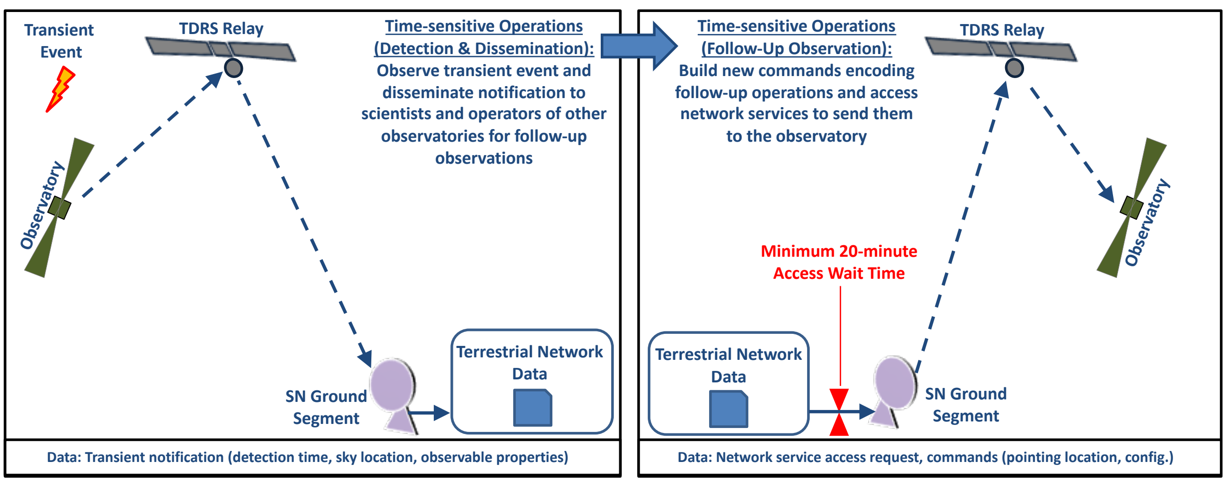
NASA's Space Network (SN)[1] provides global communications coverage for Swift and other LEO users. The SN consists of a constellation of geosynchronous Tracking and Data Relay Satellites (TDRS) and their associated ground segment. The SN Demand Access System (DAS) provides Swift with on-demand access to S-band space-to-ground (i.e., return) communications services using a code-division multiple-access scheme and pre-allocated communications

[1] A major reorganization within the Space Communications and Navigation Program occurred in October 2020. The Space Network and Near Earth Network projects were integrated into a single project office known as the Advanced Communications Capabilities for Exploration and Science Systems (ACCESS) project. Subsequent chapters in this dissertation refer to the ACCESS space relay component as the TDRS space data network.

resources [31] [32]. Upon detection of a transient gamma-ray source, Swift uses DAS to send preliminary data products, such as localization and gamma-ray flux information, to the ground within 20 seconds [31]. Additional quick-look data products associated with its X-ray and ultra-violet optical instruments may follow within minutes [31]. A transient science messaging application, known as the Gamma-ray Burst Coordinates Network (GCN), further disseminates these data products to the scientific community via the Internet using a publish-subscribe design pattern [17]. GCN information is sometimes used to initiate and coordinate follow-up observations of a transient source among scientists and operators of complimentary optical, radio or other types of space and ground-based observatories. Some ground-based observatories have implemented automated machine-to-machine interfaces to the GCN to minimize their follow-up observation latency [17]. Members of the scientific community can also submit requests for observations of transient sources to the Swift science operations team [16] [17]. Swift scientists receive approximately 1,400 such requests per year, which are categorized by urgency and priority [16]. The most time-sensitive transient sources require unplanned follow-up operations by the Swift observatory within a few minutes to less than four hours from the time of notification [16].

The SN does not presently have the capability to provide users with on-demand access to ground-to-space (i.e., forward) communications services comparable to the timely return services provided by DAS. When Swift receives an urgent and high-priority transient source notification from a GCN participating observatory, its mission operators coordinate by telephone with SN operators to manually arrange expedited access to network services. SN operators report that the minimum wait time for manually arranged service access is 20 minutes. Additionally,

Swift and SN operators report that manual intervention imposes substantial task burden. Fig. 1 below illustrates the current concept for time-sensitive multi-observatory science operations involving SN. The left panel illustrates the role of SN DAS (noted by a dashed line) in enabling transient event observation data to be returned quickly from space-based observatories for further dissemination to the scientific community via terrestrial networks. The right panel illustrates the current need for manual operator intervention (noted by the red hourglass) to request and grant network access in order to send user commands directing follow-up observations of externally identified transient events to space-based observatories.



Space observatories send unplanned data to the ground quickly using the SN Demand Access System

Operator arranged network access is current practice to send unplanned data to space observatories

Fig. 1: Current concept for distributed time-sensitive transient science operations using SN.

If timely access to bi-directional space communications services were available, it could be used to transport many types of coordination, management and control traffic among diverse users and network service providers, thus enabling a more network-centric, automated and commercialized space mission enterprise [9]. The literature provides several concepts for improving service access timeliness in direct-to-earth and space relay networks. Notably, pre-

scheduled blocks of time are reserved on NASA's Deep Space Network to monitor and respond to beacon tones signaling unplanned service needs from any of several mission spacecraft within a coverage area [33] [34]. The technique of pre-scheduling network service periods to achieve timeliness outcomes for unplanned needs is conceptually similar to the Space Network gap-filling method developed in this research. The Deep Space Network has long network configuration and inter-user setup times (45 min. and 15 min., respectively) and high service utilization, making a timely event-driven service access method impractical [34]. Previous efforts have proposed and demonstrated alternative methods for implementing timely access to SN ground-to-space network services using delay/disruption tolerant networking protocols and broadcast services [35] [36]. However, these efforts have failed to achieve stakeholder consensus on user needs and network requirements within the broader NASA mission and network services architecture. Additionally, the proposed methods have required costly engineering modifications to SN systems. The continued lack of timely access to bi-directional network services presents an obstacle to the scientific discoveries enabled by distributed time-sensitive space mission operations.

Research Strategy and Chapter Organization

This research addresses the shortcomings of prior efforts and advances the network services roadmap for enabling distributed time-sensitive space mission operations. Two methods for users to obtain timely access to network services using SN are developed within NASA's overarching mission and network architecture. The goals of the conceptual modeling effort are to establish and demonstrate a rigorous, traceable and consistent model-based systems engineering approach for describing and assessing network capabilities within the context of

driving user operational scenarios. Next, pathfinder experiments are designed within the constraints of the as-built SN systems to allow the effectiveness of the methods to be demonstrated and evaluated in the context of a time-sensitive Swift operational scenario. In addition, the pathfinder experiments are intended to deliver operational value to existing users, to increase stakeholder advocacy for investments in future network service capabilities, and to clarify system trade-offs and priorities for enabling distributed time-sensitive space operations.

Section 1 of this chapter provides background for this work and introduces the Swift user mission case study. Section 2 provides architectural context and conceptual development for a pre-determined gap-filling network access method and a random event-driven network access method, including measures of effectiveness and potential applications for each method. Section 3 presents the experimental design and results of these methods in the context of a Swift operational scenario. Section 4 provides a discussion of results and identifies next steps. Section 5 presents the conclusions from this chapter.

Conceptual Development

This section describes the conceptual development of two network service access methods using a model-based systems engineering approach and discusses applications for each method.

Architecture Context

The Space Communications and Navigation (SCaN) Program is responsible for all aspects of the space communications and navigation infrastructure used by NASA’s science and human exploration missions [37]. SCaN requirements encompass electromagnetic spectrum policy

coordination, advanced technology development and operational network services for user platforms located on or near the Earth, other heavenly bodies and beyond the solar system. SCaN has identified eight broad groupings of use cases, which identify the primary network behaviors and attributes involved in meeting the requirements of its diverse users. SCaN has also defined a set of user need categories and associated parameters which provide a basis for assessing the effectiveness and efficiency of its network capabilities and services across the use case groups. SCaN user need categories are defined and described in Table 4.

Table 4: Network user need categories.

User Need Categories	Description
Data Volume	The amount of data generated or received by the user within a defined period of time. Includes science mission data, engineering telemetry, commands and navigation data.
Latency	The allowable time to deliver data from a source to a destination. Mission latency needs may vary by data type, application or operational scenario.
Dependability	The availability, proficiency and reliability of network services for which network service providers are held accountable to users.
Access	The ability to obtain network services at a specified time, or with a specified timeliness or with a specified frequency.
Navigation	The ability to perform orbit determination, maneuvering and navigation functions based on accurate and precise network time and frequency references, signal observables or data services.
Mission and Platform Specific	The ability of network solutions to satisfy user mission and platform needs and constraints (e.g. operational constraints imposed by science activities, orbital eclipse periods, platform size, weight, power etc.)

This research is concerned with improving the effectiveness and efficiency of network access in support of transient science user operations that involve coordinated and time-sensitive follow-up activities across multiple observatories. Such operations are associated with two

groups of SCaN use cases: 1) near-Earth robotic users with low latency and complex data transport needs and 2) user and network mission operations.

Several Systems Modeling Language (SysML) diagrams were developed iteratively to relate and further specify the system architecture associated with the two applicable use case groups. SysML provides a standardized visual vocabulary for describing system architectures using a set of interconnected object-oriented diagrams [24]. SysML is useful for ensuring traceability, consistency, and rigorous specification of system representations at all levels in the system hierarchy [24]. A system-level block definition diagram and activity diagram were developed to demonstrate traceability and conformance to the SCaN architecture. These diagrams serve as the root of the gap-filling and event-driven access methods developed in this research.

A system block definition diagram represents structural aspects of the system, including its boundary, hierarchy and primary interfaces, using blocks that group related functions and resources [24]. The structural model of a transient science space system is presented as a SysML block definition diagram in Fig. 2.

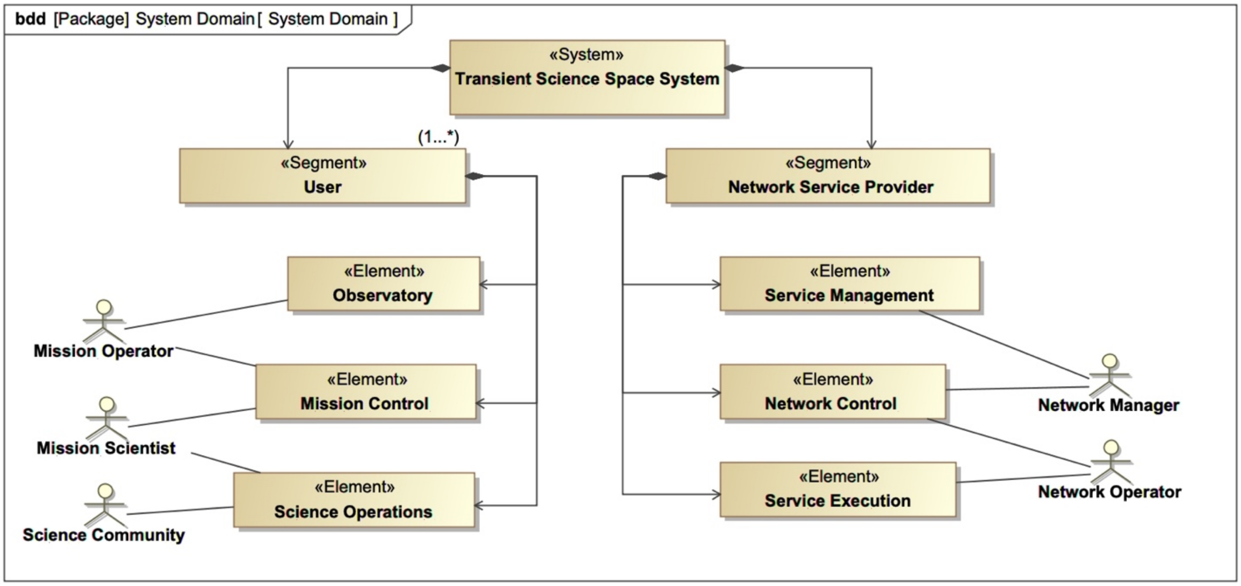


Fig. 2: Structural model of a transient science space system.

A transient science space system is comprised of one or more user segments which interact collaboratively and dynamically through a network service provider segment on timescales that may range from tens of seconds up to several hours. The user segment is further decomposed into three elements supported by three actor classes. The observatory element performs scientific measurements, processes and stores data and commands, and executes spaceflight functions. The mission control element supports the management and control of the technical aspects of observatory operations, including specialized components to create and verify command loads for observatory flight software and to build network service access requests. The mission operator actor uses the observatory and mission control elements to perform user operations. The science operations element supports the management and control of the scientific aspects of observatory operations, including scientific target priorities and scientific data analysis. The mission scientist actor uses the science operations and mission control elements to fulfill the mission’s scientific objectives. The science community publishes

notifications of transient scientific events associated with the science operations of external space or ground-based observatories to terrestrial Internet Protocol networks. When notified of a transient scientific event from the science community, the mission scientist must decide whether to preempt planned activities in order to conduct follow-up observations. The mission scientist works closely with the mission operator to implement follow-up operations since new observatory commands specifying the observation parameters must be built and unplanned access to network services may be necessary to deliver the commands to the observatory within the overall latency constraints of the observation scenario. The network service provider segment is partitioned in conformance to the interfaces established in the SCaN Architecture Definition Document. It is comprised of three elements and two actors. The abstract network manager actor class represents several specialized roles, including mission commitment managers and service scheduling planners, among others. For the purposes of this research, this actor class is responsible for interacting with users about their network service delivery needs occurring more than one week into the future. Network operators are responsible for the operational monitoring and control functions of the network service provider, including interactions with users for unplanned service access. The service management element ingests service access requests from users. This information is used by network managers to develop deconflicted and prioritized service schedules among all users during the schedule planning phase. The service management interface is also used to provide service access request status information to users. The network control element ingests valid network service tasks provided by the service management element or manually by network operators, issues configuration directives to network resources, and reports status to network operators. Finally, the service

execution element performs the functions associated with establishing an end-to-end data transport path between the user's terrestrial and space-based element (i.e., the mission control and observatory elements).

A system activity diagram represents behavioral aspects of the system by defining and relating a set of activities and information involved in performing an operational workflow [16]. The activities are allocated to the structural elements defined in the system block definition diagram, ensuring cohesion of the overall conceptual model [16]. The behavioral model for a transient science space system conducting follow-up observations is presented as a SysML activity diagram in Fig. 3. This representation places the contribution of network access timeliness within the context of other system latency contributors within the transient science follow-up observations operations concept.

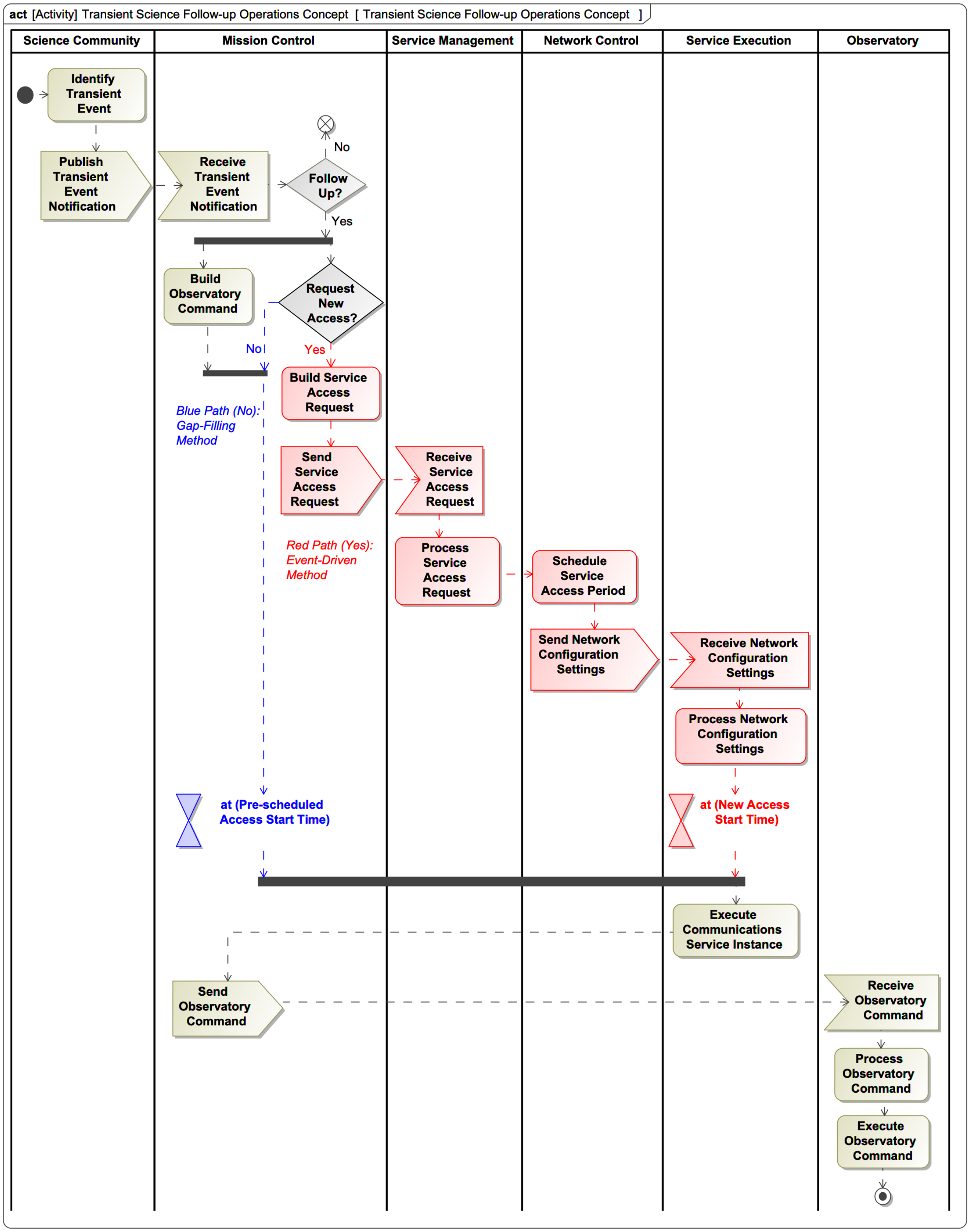


Fig. 3: Operations concept for follow-up observations of an externally identified transient scientific source.

The following preconditions apply to this operations concept: the user mission is conducting nominal pre-planned operations and has been granted network service access periods through the pre-determined network scheduling process. The activity flow is initiated by detection of a transient scientific event by the external science community. The science community identifies salient parameters characterizing the transient event, such as its celestial location and time of detection, and publishes this data as a notification to subscribed space mission users over terrestrial Internet Protocol networks. The transient notification is received by the mission control element. The mission scientist must decide (or delegate to automation) whether to preempt planned activities in order to conduct follow-up operations. This choice is represented by the first decision node in the activity flow (labeled "Follow Up?"). If the mission scientist elects to perform follow-up operations, then a new flight software command load specifying the desired observatory operations must be built by the mission operator or generated autonomously by the mission control element. Timely network service access is needed to send the new command load to the observatory. As a result, the mission scientist must decide (or delegate to automation) whether the next pre-scheduled service access period is sufficiently soon to meet the overall latency objective for the follow-up observation, or if more timely access is necessary. This choice is represented by the second decision node in the activity flow (labeled "Request New Access?"). The two methods developed in this research are pertinent to this second decision. As will be developed further in subsequent sections, the gap-filling method achieves guaranteed timeliness outcomes by increasing the frequency of pre-scheduled service periods during direct-to-earth network coverage gaps, while the event-driven method achieves best-effort timeliness outcomes by reserving a new service period during real-time operations

using novel rapid access procedures. The gap-filling method and event-driven method are illustrated respectively by the blue and red activity flow branches. Since the gap-filling method is implemented during the nominal schedule pre-planning phase, if the mission scientist decides not to request new service access, then the new command load must wait to be sent until the next pre-scheduled service access start time, as represented by the blue hourglass. If the mission scientist decides to request new service access, then a sequence of activities to request, reserve and provision services are performed by the system elements indicated in the diagram. Each of these activities contribute to the expected wait time of the event-driven method, represented by the red hourglass. Network services are executed at the service access start time, allowing the new command load to be sent from the mission control element to the observatory element. The observatory element then receives, processes and executes the commands. As a postcondition to this operations concept, the observatory completes the follow-up operations specified by the new command load.

It is important to note that the conceptual model of the transient science space system is abstract. In a fully implemented transient science space system, use of either or both methods may be suitable. Further specification of the transient science space system model is necessary to assess the effectiveness and efficiency of the methods for their application in a defined operational scenario.

Gap-Filling Network Access Method Definition and Development

At the present time, NASA's network operations are planned two or more weeks in advance of service execution. Mission users forecast their service needs, compute the set of visibility periods between their orbiting spacecraft and compatible ground or relay

communications networks, and formulate service requests in terms of a service configuration code and access time on specific communications nodes. Service providers evaluate these requests based on pre-established priority lists and negotiate conflicts among users. A de-conflicted service schedule is disseminated to users approximately one week before the start of the active execution period. The active schedule execution period is also referred to as the real-time operations phase. The schedule defines a fixed batch of communications service periods over seven days. Prior to the start of the active period, users and network actors build and load command sequences to their respective elements that will execute the specified configurations according to the schedule. Under nominal conditions, the user and network system elements execute the schedule autonomously, with varying degrees of supervision by mission and network operators.

Swift's concept of operations relies on obtaining access to services from two complimentary communications networks. The Swift observatory has two onboard communications subsystems, each tailored for its purpose and for ground or relay network compatibility. Service access requests for the ground network and space relay network (i.e. SN) are handled separately, with different user interfaces and service specification parameters. Swift's typical ground network schedule provides communications access roughly every 30 to 120 minutes, with larger service gaps of 4 to 8 hours also possible due to variations in orbit coverage and resource blocking by higher-priority users. SN DAS resources are pre-allocated so as to be continuously available as the observatory transits the SN's Atlantic, Pacific and Indian Ocean regional service areas. Timely on-demand access is achieved by configuring the Swift relay communications subsystem with pointing and handover information for a reserved SN DAS

resource. Although the Swift observatory is compatible with SN pre-scheduled S-band multiple-access services, these capabilities are not routinely used. However, for spacecraft emergencies and the highest-priority transient event notifications, SN service access is requested and allocated manually.

Many transient events of scientific interest occur following a Poisson random arrival process [15] [38] [3]. As a result, it is possible to statistically analyze the access timeliness characteristics for fixed service schedules [39]. The access timeliness characteristics for a fixed schedule can then be evaluated for suitability for a given time-sensitive operational scenario. Three measures of effectiveness are identified and applied from the literature to characterize access timeliness for fixed service schedules: the maximum gap duration, Gap_{Max}, and its likelihood of occurrence $P[Gap_{Max}]$, the expected (mean) gap duration, $E[Gap]$ and the expected user wait time to access services, $E[Access]$ [20].

Consistent with the abstract concept of operations defined previously in Fig. 3, a time-sensitive operational scenario is triggered by the random observation and notification of a transient event by the science community. Transient notifications may follow the arrival rate distribution characteristics of the scientific observable or some other distribution depending on the details of a given operational scenario. The arrival rate distribution can be used to predict the number of occurrences expected within a given fixed schedule (or any subset interval) [39]. This information is useful in evaluating the feasibility and suitability of a given gap-filling method implementation. Regardless of the transient notification arrival rate, the specific arrival time of a given notification is completely random within a fixed schedule interval and is therefore equally likely to occur at any moment in time within it [39].

Continuous access to return communications services, such as with SN DAS, allows notifications of transient events detected by space-based observatories to be disseminated on-demand. However, when external (e.g., ground or other space-based) observatories generate a transient notification, bi-directional communications for sending and verifying receipt of commands to a space-based observatory must wait for the next pre-scheduled communications event to occur[2]. The probability that a transient notification will arrive in an interval where there is a service gap can be computed as the sum of gap durations (Gap_i) divided by the duration of the fixed schedule ($T_f - T_0$). The service gap of maximum duration, Gap_{Max}, is also the most probable gap within which a transient notification should be expected to arrive (i.e., Gap_{Max} is the mode of the probability distribution of service gaps within a fixed schedule). The likelihood that a transient notification will arrive within the maximum gap duration of a fixed schedule is given by Eq. 1.

$$P[Gap_{Max}] = \frac{Gap_{Max}}{\left(T_f - T_0\right)} \tag{1}$$

[2] In principle, if a transient notification arrives during a service period, then the access wait time is zero. However, in practice it may be difficult for current systems to process a transient event notification, generate new commands that update or replace the planned command set for the service period, send the new commands to the observatory and verify their receipt within the remaining duration (i.e., following the notification arrival) of a typical 8 to 10 minute ground network service period. As a result, a more sophisticated model than presented in this chapter could be developed that would account for the response processing latencies (i.e., the sum of the factors identified above), the probability of usable service periods given a notification arrival during a service period (i.e., those that contribute zero access wait time), and the increased wait time that would result from unusable service periods. The magnitude of the wait time contribution due to unusable service periods could be bounded by the sum of the response processing time, the service period duration, and the duration of the subsequent service gap. However, the probability of a transient notification arriving during a service period is typically much less than the probability of an arrival during a gap (roughly 10% and 90%, respectively), and unusable service periods would occur at some fraction of the small service period arrival probability. As a result, the overall impact of accounting for these complicating factors on the total expected access wait time for the gap-filling method is small. Accordingly, the model presented in this chapter provides a simplified but reasonable estimate of the expected access wait time.

The expected service gap, *E[Gap]*, is the mathematical mean of the distribution of discrete gap durations, Gap_i, present in a fixed service schedule interval $(T_f - T_0)$. It measures the average length of the service gap found if the gap durations are sampled randomly, as is the case with transient notification arrivals. It is constructed as the sum of each gap duration multiplied by its probability of occurrence within the fixed schedule interval, given by Eq. 2.

$$E[Gap] = \sum_{i=1}^{n} Gap_i \left(\frac{Gap_i}{(T_f - T_0)} \right) \tag{2}$$

Random transient notifications are equally likely to arrive at any moment within the interval of the mean gap duration. For example, if a transient notification arrives at the first instant of a service gap, the wait time to service access is approximately equal to the gap duration. However, if the transient notification arrives at the last instant of a service gap, the wait time is approximately zero. Over a large number of arrivals, the average wait time within the gap interval approaches one half of the expected service gap duration, given by Eq. 3.[3]

$$E[Access_{GapFilling}] = \frac{E[Gap]}{2} \tag{3}$$

Fig. 4 provides a model and summary of the access timeliness measures of effectiveness for fixed service access schedules, illustrated with a notional schedule and three random transient notification arrival events.

[3] A formal proof is presented in Tijms [19, pp. 15-17]

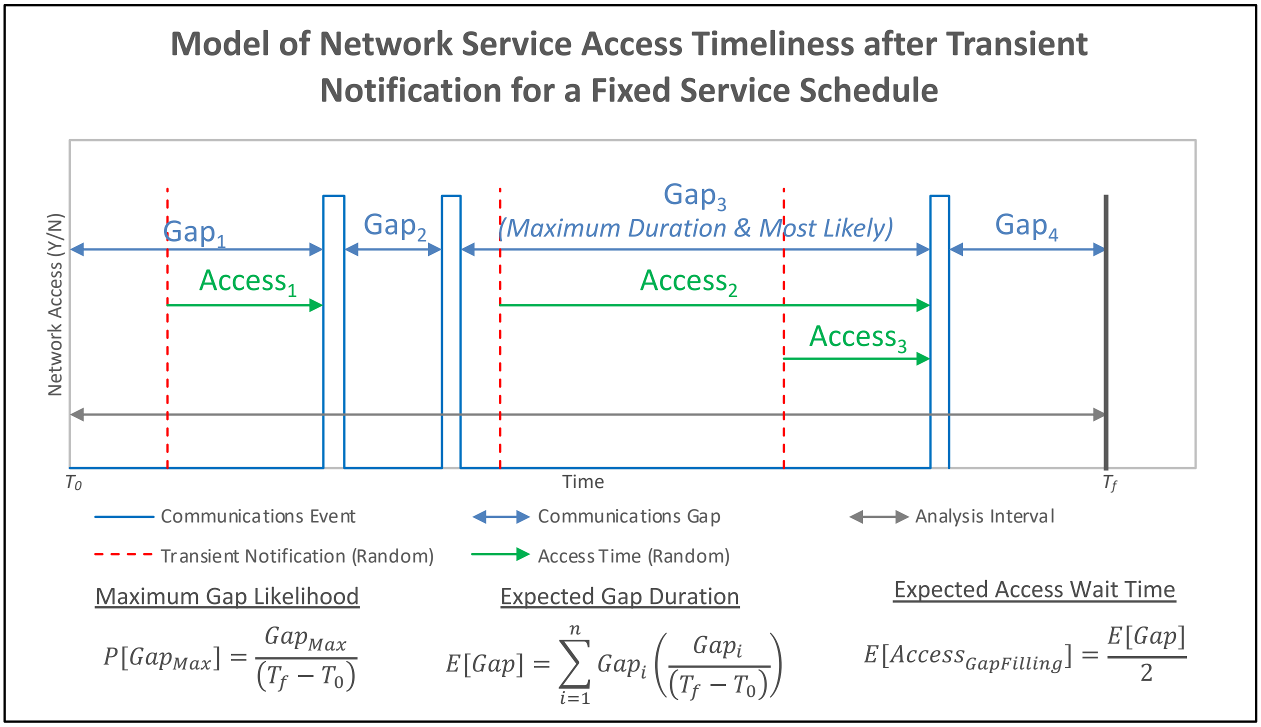


Fig. 4: Model of network service access timeliness for a fixed service schedule.

For missions that only have access to a single ground network service provider, the access timeliness measures of effectiveness are governed by the constraints imposed by coverage, orbital geometry and the network loading from other users. However, missions that have access to more than one network provider (including space relay networks or other ground networks with complimentary coverage) may influence the access timeliness measures of effectiveness for fixed schedules by requesting secondary network service periods to fill the largest and most probable primary network service gaps during the nominal schedule pre-planning phase. In this way, the gap-filling method relies on the frequency of fixed schedule service periods from all provider networks to achieve desired access timeliness outcomes. The gap-filling method is illustrated by the blue colored branch of the concept of operations defined previously in Fig. 3. In the case of Swift, the SN could serve as such a secondary network provider. Since the combined

primary and secondary network schedules are pre-determined, the gap-filling method can be used to ensure access timeliness outcomes without the need for real-time manual intervention to access services. The disadvantages include the expenditure of planning resources to reserve secondary network service periods that may not be used (i.e., if a transient notification does not arrive) and potential opportunity costs of the reserved secondary network resources for other uses. Missions seeking to implement the gap-filling method should account for the transient notification arrival rate as well as the temporal decay rate of the scientific value associated with the transient observable in the time-sensitive operational scenario. These factors should be compared to the feasibly achievable access timeliness outcomes of a gap-filling method implementation, with a full accounting of costs and benefits among mission and network stakeholders. For example, although a hypothetical long-term gap-filling method implementation for a rarely occurring transient observable with fast-decaying scientific value may impose prohibitive stakeholder costs, perhaps stakeholders might agree to implement a short-term campaign in order to demonstrate or validate the potential value of a new multi-observatory workflow or to gather new observable signatures for development and training of future onboard algorithms. A notional representation of the gap-filling method concept of operations is provided in Fig. 5 below.

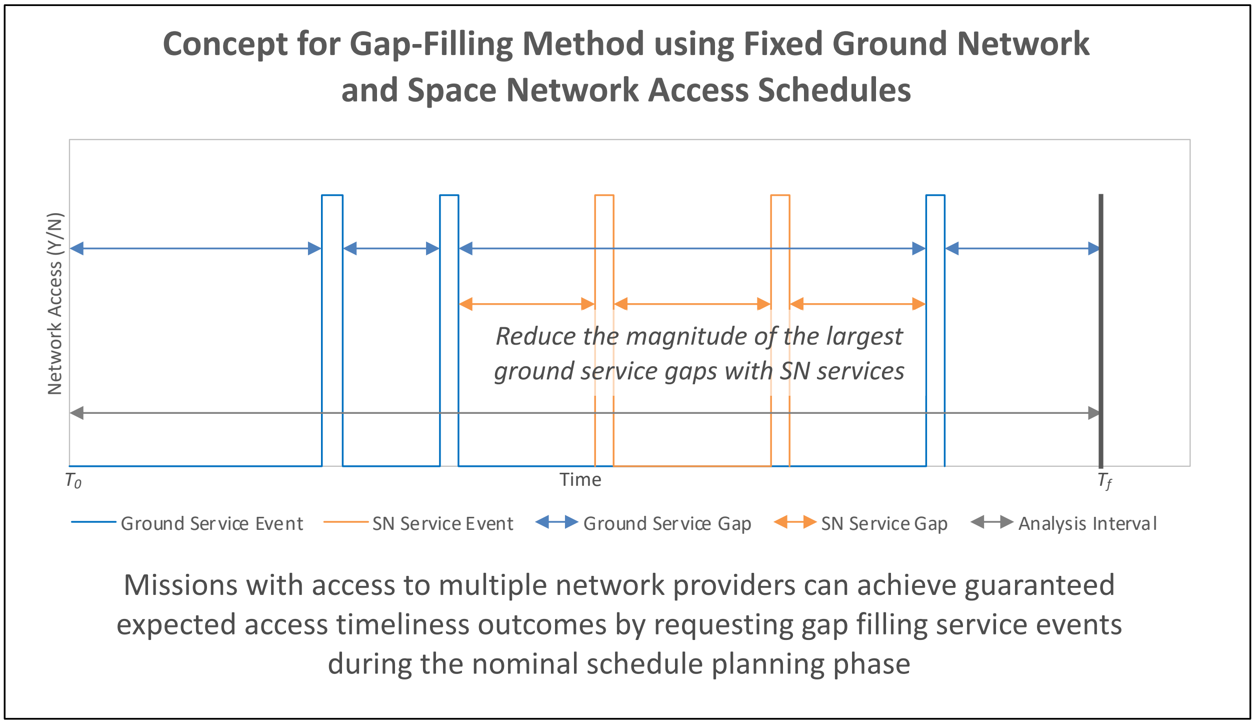


Fig. 5: Concept for achieving guaranteed expected access timeliness outcomes using SN as a secondary provider.

Event-Driven Network Access Method Definition and Development

The event-driven network access method involves the activities indicated by the red colored branch of the concept of operations defined previously in Fig. 3. There are several pre-conditions that must be satisfied in order to implement an event-driven service access method. First, it should be possible to close the free-space communications link for the greatest possible duration over the observatory's trajectory, as unavoidable service gaps occur when the link cannot be closed. This is influenced by observatory and network factors such as transmit power, receiver sensitivity, antenna gain patterns, error correction coding, visibility and separation distance, among others. Second, since the service instances are unplanned, the observatory must be in an "always listening" state to detect and acquire the initiation of services. To satisfy the first two preconditions, the observatory requires precise timing and navigational-state awareness, a

pre-defined set of possible service access nodes for antenna pointing and tracking, and the ability to correct for trajectory induced doppler frequency effects. The Swift observatory satisfies these preconditions for SN services. Next, network systems must be able to accept and disposition service access requests during the active schedule execution period. Swift's most time-sensitive transient notifications require service access within a few minutes up to four hours. To meet these demands, network systems must also have the ability to execute service periods within minutes or hours following receipt of the access request. This requires sufficiently available resource capacity or advanced traffic management capabilities. Over the summer of 2019, the authors held a series of discussions with SN operators, conducted a thorough review of SN documentation and performed a preliminary analysis of a resource availability dataset for the TDRS multiple-access S-band services. Results from this work established that SN systems satisfy the preconditions for an event-driven service access method. Additionally, the authors identified latent, but little known or used SN capabilities that allow users to incorporate flexibility in the specification of service requests, which can result in improved access wait times. The contributors to service access wait time outcomes for the SN event-driven method are identified and developed subsequently.

For nominal pre-planned operations, SN users build and submit batches of service access requests during the fixed schedule planning phase using the SN's service management interface to the MOC. This allows users to request and reserve future service periods using a standard set of messages and configuration parameters that are exchanged between the user MOC systems and SN systems. Once granted, SN service periods are guaranteed in the weekly fixed schedule, except in rare instances of unplanned outages or spacecraft emergencies.

The SN event-driven service access method occurs within the active period of fixed SN schedules and relies on the same service management interface used for fixed schedule service planning. Users may build and submit service access request messages during the active period provided they conform to a set of scheduling ground rules. One such rule specifies that the minimum lead time between receipt of a service access request message and the start of the service period shall be no less than 10 minutes for all SN service types. This lead time accounts for the request processing and SN system configuration setup time. The largest wait time contributor to the 10-minute service start lead time is the 6-minute worst-case inter-user pointing time for the TDRS mechanically steered antenna used for SN single-access services. However, the SN event-driven method uses multiple-access services, which rely on the TDRS electronically steered phased array resources. The inter-user setup time for multiple-access services is specified as 30 seconds. This suggests that the minimum service lead time may be significantly reduced if, in the future, the SN systems could filter requests for multiple-access services from requests for single-access services and apply differing processing constraints.

An important measure of effectiveness for time-sensitive operations is the minimum service access wait time, $Access_{Min}$. The minimum service access wait time is achievable when the requested SN resource is available and depends on the wait time contributions of the service access request message build process (*RequestBuild*) and the specified minimum service start lead time (*MinimumLead*). The minimum service access wait time is given by Eq. 4.

$$Access_{Min} = RequestBuild + MinimumLead \tag{4}$$

Unlike the guaranteed access provided in fixed schedules, event-driven service access requests are fulfilled using SN's unreserved resource capacity on a first-come-first-served basis. Event-driven service access is blocked by previously reserved time periods for user services and SN maintenance and sustainment activities. As a result, it is important to characterize the durations of blocking periods, $Blocking_i$, for the event-driven method. SN publishes a registry of available resources during the active period, known as the TDRS Unused Time (TUT) registry. The TUT registry includes available and reserved time periods for each TDRS node and service type. The set of blocking periods is measured directly from the TUT registry as the reserved intervals. This allows calculation of the SN blocking likelihood, *P[Blocking]*, and the expected duration of blocking periods *E[Blocking]*.

The blocking probability is computed as the sum of blocking period durations divided by the fixed schedule analysis interval ($T_f - T_0$), described in Eq. 5.[4]

$$P[Blocking] = \sum_{i=1}^{n} \left(\frac{Blocking_i}{(T_f - T_0)} \right) \tag{5}$$

The expected blocking duration, *E[Blocking]*, is the mathematical mean of the distribution of blocking durations present over the analysis interval. It measures the average duration of the blocking duration found if the blocking durations are sampled randomly, similar to the approach employed in Eq. 2. It is constructed as the sum of each blocking duration, $Blocking_i$, multiplied by its probability of occurrence, *P[Blocking]*, given by Eq. 6.

[4] Note, the probability that a TDRS resource is available can be computed similarly, but because resources must either be reserved or available, the probability of resource availability is also equal to one minus the blocking probability.

$$E[Blocking] = \sum_{i=1}^{n} Blocking_i \left(\frac{Blocking_i}{(T_f - T_0)} \right) \quad (6)$$

The magnitude of the expected blocking duration can be reduced by increasing the number of simultaneous multiple-access forward service instances available in an SN service region. This could be achieved by adding additional TDRS nodes in a region or by splitting the forward array on second and third generation TDRS nodes into two independent beams. However, this would require modification to the as-built SN systems.

The expected service access wait time if an SN resource is blocked, $E[Access_{Blocking}]$, encompasses the two wait time contributors found in the minimum access wait time calculation ($Access_{Min}$, provided by Eq. (4) as well as the added wait time contributed by the expected blocking duration. SN multiple-access resources must also be reconfigured between uses. As a result, a final wait time contributor, (*InterUserSetup*), occurs following an SN blocking period. The expected service access wait time, $E[Access_{Blocking}]$, if blocking occurs is provided in Eq. 7.

$$E\left[Access_{Blocking}\right] = Access_{Min} + \frac{E[Blocking]}{2} + InterUserSetup \quad (7)$$

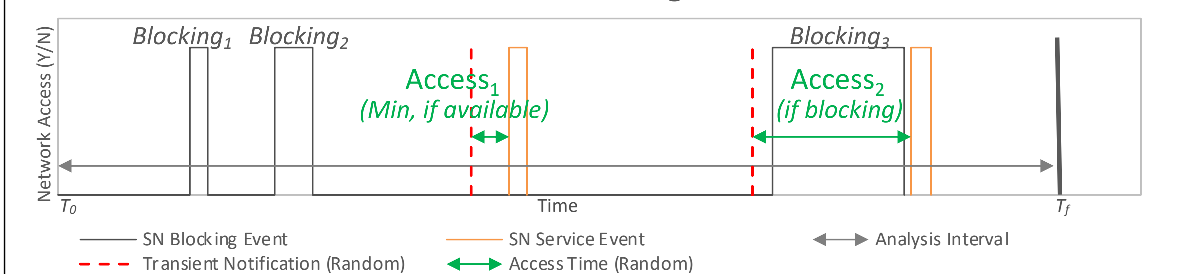
Finally, the total expected service access wait time for the SN event driven method can be calculated as a probability weighted sum of the access wait time outcomes for when a resource is available ($Access_{Min}$) or blocked ($E[Access_{Blocking}]$), provided by Eq. 4 and Eq. 7 respectively. The total expected service access wait time for the event-driven method, $E[Access_{EventDriven}]$, is given by Eq. 8.

$$E[Access_{EventDriven}] = P[Blocking] * E\left[Access_{Blocking}\right] + (1 - P[Blocking]) * Access_{Min} \quad (8)$$

The measures of effectiveness for the event-driven method can be used by mission and network planners to assess the feasibility and value of time-sensitive mission operations concepts. First, the minimum access wait time is useful as a screening mechanism to filter out time-sensitive concepts which require more timely service access than is achievable with SN systems or to identify the degree to which SN wait time contributors must be improved to make a mission concept viable. The probability of blocking provides users with information about how frequently the minimum and blocking access wait time outcomes will occur. The expected duration of blocking periods is useful to users for the specification of flexible service start times. As part of the Schedule Add Request message, SN allows users to indicate service start time flexibility as a plus or minus time tolerance. Service start time flexibility reduces the likelihood that a requested service instance will be blocked. The SN scheduling system will shift the requested service start time according to the specified start time tolerances, ensuring users achieve the minimum wait time by allocating a service instance immediately preceding or following a blocking period. The expected duration of blocking period is also a useful metric for network managers since its magnitude can be reduced by increasing the capacity of supportable service instances within a given service region. The expected access wait time when blocking occurs provides the average wait time outcome when resources are not available. This metric can be used by mission planners to judge the expected value of a time-sensitive operations concept when resources are not available. Finally, the total expected service access wait time for the SN event-driven method provides a weighted average metric for the event-driven method. Users conducting time-sensitive operations are likely to have non-linear timeliness preferences, with their value of an outcome decaying as a function of time. As a result, users may more clearly

judge the value of service access timeliness based on the component outcomes and probabilities (i.e., wait times experienced if SN resources are available or blocked) rather than the more abstract probability weighted average, which does not represent a realized operational outcome.

Fig. 6 provides a model of the SN event-driven method and a summary of the measures of effectiveness that are most salient for user access wait time outcomes with the event-driven method. It is comprised of a notional set of blocking periods and two event-driven service instances illustrating access wait times when resources are available and when blocked. Equations for the highlighted measures of effectiveness are also represented.



Likelihood of SN Resource Availability

$$P[Available] = (1 - P[Blocking])$$

Likelihood of SN Resource Blocking

$$P[Blocking] = \sum_{i=1}^{n} \left(\frac{Blocking_i}{(T_f - T_0)} \right)$$

Minimum Access Wait Time (if Available)

$$Access_{Min} = RequestBuild + MinimumLead$$

Expected Access Wait Time (if Blocking)

$$E[Access_{Blocking}] = Access_{Min} + \frac{E[Blocking]}{2} + InterUserSetup$$

Fig. 6: Model of network service access timeliness for an SN event-driven service access method.

The event-driven access method allows users to request service access as needed, reducing the operational planning and network utilization costs imposed by the gap-filling method. However, service access is not guaranteed due to resource blocking by previously reserved events. Missions seeking to implement the event-driven method should account for the

transient notification arrival rate as well as the temporal decay rate of the scientific value associated with the transient observable in the time-sensitive operational scenario. These factors should be compared to the achievable access wait time outcomes provided by the SN event-driven method, with a full accounting of costs and benefits among mission and network stakeholders. Since the event-driven method involves claiming small windows of available resource time for near-term use, it is suitable for long-term transient observation campaigns or campaigns with rarely occurring transient observables. Such campaigns would likely impose prohibitive network utilization costs if implemented using the gap-filling method. Applications for the event-driven method may include demonstration or validation of a new multi-observatory operations workflow or gathering new observable signatures for development and training of future onboard observatory algorithms.

Experimental Design & Results

This section develops the service access method pathfinder experiments. It provides details for a specific time-sensitive mission operational scenario, discusses the experimental setup and application of the two service access methods in the context of this scenario, and presents the experimental results.

Swift Operational Scenario Definition

A Swift follow-up observation operational scenario was chosen to demonstrate and investigate the suitability of the two service access methods. This operational scenario is consistent with and traceable to the abstract operations concept presented in Fig. 3 and provides the additional details necessary to assess the effectiveness and efficiency of the methods.

Detecting the electromagnetic counterparts associated with gravitational waves is a major new scientific objective in astrophysics. Detectable gravitational-waves are observed with ground-based observatories with an expected arrival rate of one per week [40]. Swift is capable of observing the gamma-ray, X-ray and ultra-violet optical electromagnetic counterparts to gravitational-waves. However, the Swift observatory cannot itself detect gravitational-waves or identify which electromagnetic transient events are associated with them. The Swift observatory continuously feeds instrument data into a working buffer where the data is processed to identify and classify observed transient events. Electromagnetic events which satisfy an observable threshold result in the creation of a transient notification, which is disseminated to the ground through the SN DAS. In addition, the high-fidelity raw data associated with the observation is transferred from the working buffer into persistent storage for downlink to the ground network at the next pre-scheduled access time. Swift's working data buffer is overwritten every 30 minutes. When a ground-based observatory detects a gravitational-wave, a transient notification is created and sent to the Swift MOC via the GCN terrestrial network. If any electromagnetic counterparts were coincidentally observed by the Swift observatory, but do not satisfy the onboard observable threshold, this data is at risk of being overwritten in the working buffer and lost. Therefore, whenever the Swift MOC receives a transient notification from a gravitational wave observatory on the GCN indicating a gravitational wave detection, a follow-up command to transfer data from the observatory's working buffer to its persistent storage must be executed by the observatory within 30 minutes of the gravitational wave detection event.

Gap-Filling Network Access Method Experimental Design and Results

In August 2018, Swift and SN operators held a teleconference to plan a gap-filling method pathfinder experiment. The primary objective of the pathfinder was to measure and assess the benefits and costs to Swift and SN stakeholders to determine if the gap-filling method was truly implementable in an operational environment. As a preliminary strategy, meeting participants agreed that an SN service period would be requested approximately every 60 minutes during the largest ground network coverage gaps, targeting an expected access wait time of 30 minutes. The Swift operations team agreed to provide forecasted view periods between the Swift observatory and TDRS nodes to the SN scheduling office no later than 12 days before the start of the active period week. Swift operators requested a minimum service period duration of 10 minutes to ensure adequate time for signal acquisition and for the Swift flight and ground systems to connect and synchronize. Swift accepted a "low" SN priority designation for the adjudication of gap-filling service periods.

On 2018 Day-of-Year (DOY) 304 Swift operators submitted a batch of nine SN service access requests during the largest anticipated ground network coverage gap for DOY 316. Swift operators received confirmation of both ground network and SN fixed schedules by DOY 310. Six SN service access requests were granted, resulting in an access request success rate of 66.7%. The granted ground network schedule had a maximum gap duration of 4:03:34 (H:MM:SS), beginning at 6:04:46 UTC and ending at 10:08:20 UTC. The next largest gap of 2:56:25 (H:MM:SS) began at 10:21:08 UTC and lasted until 13:17:33 UTC. No attempt was made to reduce the third largest gap of 2:42:58 (H:MM:SS), from 13:26:48 UTC to 16:09:46 UTC. No data were provided for ground network contacts before or after DOY 316. As a result, analysis interval for the

experiment is defined from the first and last known ground service periods, from 3:12:29 UTC to 19:32:48 UTC. With a gravitational wave arrival rate of one per week, Swift's probability of needing to conduct a gravitational wave follow-up operation within the experimental period was 9.6%.

Notably, Swift operators made duplicative requests for service periods beginning at approximately 6:00:00 UTC and 10:30:00 UTC. This was done because of the uncertainty that requested service periods would be granted. Swift's true need was to have either a ground network or SN service period to avoid a potential gap greater than eight hours beginning at 4:53:57 UTC to 13:17:33 UTC. The duplicative 6:00:00 service periods were both granted, yet only the ground network service period was granted at approximately 10:30:00 UTC. A summary of Swift's requested and granted communications service periods for DOY 316 is provided in Fig. 7 below.

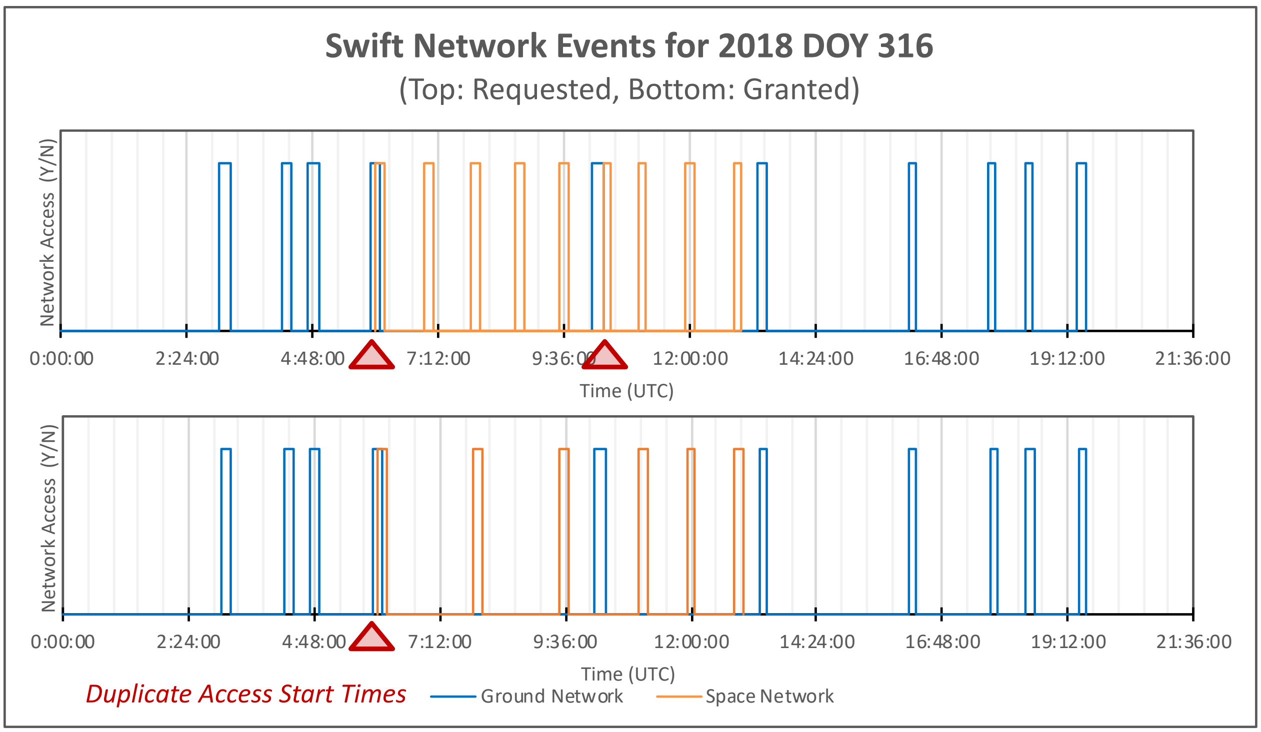


Fig. 7: Summary of requested and granted gap-filling service periods for 2018 DOY 316.

Table 5 provides a summary of the measures of effectiveness for the pathfinder experiment considering three fixed schedule cases: ground network only, ground network and the requested SN service periods, and ground network and the granted (actual) SN service periods.

Table 5: Summary of Swift's access timeliness outcomes from the gap-filling experimental pathfinder.

Fixed Schedule Cases	**Maximum Gap Duration, Gap_{Max}, & Likelihood, $P[Gap_{Max}]$ (H:MM:SS, %)**	**Expected Gap Duration, $E[Gap]$ (H:MM:SS)**	**Expected Service Access Wait Time, $E[Access_{GapFilling}]$ (H:MM:SS)**
Ground Network Only	4:03:34, 25%	2:19:19	1:09:39
Ground + Space Network (Requested)	2:42:58, 17%	0:59:56	0:29:58
Ground + Space Network (Granted)	2:42:58, 17%	1:11:44	0:35:52

Although no gravitational wave notifications arrived during the experimental period, augmenting the ground network schedule with pre-planned SN service periods resulted in a 49% reduction in expected service access wait time as compared to the ground network only case. The expected access wait time for the granted service periods exceeded the 30-minute threshold for Swift’s gravitational wave follow-up operational scenario. However, the timeliness threshold would have been met if the SN had adequate capacity to grant all of the requested service periods. In addition, the experimental results compare favorably to the minimum 20-minute access wait time achievable with the state-of-the-practice manual intervention method.

The Swift operations team reported 1.0 additional hour of effort for the planning of the nine SN service access requests. The achieved reduction in access time satisfies many of Swift’s other time-sensitive operational scenarios which require responses in less than four hours.

Importantly, the timeliness improvements were achieved without the need for real-time operator intervention.

Event-Driven Network Access Method Experimental Design and Results

The authors installed an instance of the Space Network's service management element, known as the Space Network Access System (SNAS), at the Goddard Space Flight Center's Communications Standards and Technology Laboratory (CSTL) in October 2019. The CSTL SNAS software client was configured to replicate the Swift MOC SNAS instance and connected via terrestrial mission network to the SN's Alternate Network Control Center (ANCC) in White Sands, New Mexico. The ANCC is the high-fidelity engineering test instance of the SN's network control element. Together these systems comprise the experimental setup for the event-driven service access method.

The first step for users of the event-driven method to access SN services is to build an SN service access request (*RequestBuild*). The specific message type for SN users to request service access is known as a Schedule Add Request (SAR). Users have three options, with varying degrees of autonomy, for building and submitting SAR messages. The most basic option to create a SAR involves parameter specification via text fields and pull-down menus through the SNAS graphical interface. According to SN scheduling ground rules, the specified service period start time parameter must be no sooner than 10 minutes following receipt of the SAR by the SN network control element. To build a request resulting in the minimum achievable service access wait time, users must compute a time offset following receipt of the transient notification that includes the request build time and the minimum lead time to service period start.

An experimental service access request build procedure which results in a SAR message specifying the minimum allowable wait time (*$Access_{Min}$*) was documented for the SNAS graphical interface. The authors executed the procedure several times to benchmark SAR build times. Valid and fully specified SAR messages were consistently built in the SNAS user interface within 40 seconds without haste. The procedure was provided to Swift mission operators in December 2019. SNAS also allows users to simplify scheduling of repeated service instances that have the same structure by defining "prototype" SAR messages. Prototype SAR messages reduce the build time and potential for user input error by allowing users to save the configuration of parameters that do not change. All parameters specifying a service instance with the minimum service access wait time, except for the service start time offset, can be saved as a prototype SAR message. As a result, use of the prototype SAR message can reduce SAR build times to approximately 15 seconds. Finally, SN users may implement an External Processing System to extend and automate the service management capabilities of SNAS. Accordingly, computation of the minimum service start time offset and complete specification of an event-driven SAR message can be fully automated. A summary of the SAR build times (*RequestBuild*) for each build option is presented in Table 6 below.

Table 6: Summary of SN service access request build options and associated build times.

Schedule Add Request (SAR) Message Build Options	**Build Time, *RequestBuild* (H:MM:SS)**	**Basis**
SNAS Graphical User Interface	0:00:40	Demonstration
Prototype Schedule Add Request	0:00:15	Demonstration
External Processing System (Fully Automated)	0:00:05	Estimate

The authors conducted verification testing of the documented scheduling ground rules that contribute to access wait time using the CSTL SNAS instance and the ANCC. Limited test results indicated that the documented 10-minute minimum lead time (*MinimumLead*) can be reduced to six minutes plus a variable SAR processing delay ranging from 2 to 23 seconds. However, this reduced minimum lead time has not yet been validated on the operational SN network control element. Accordingly, the baseline scheduling ground rule of a 10-minute minimum lead time is presented for the results in this chapter. The documented multiple-access service scheduling ground rule for inter-user setup time (*InterUserSetup*) of 30 seconds was experimentally verified.

The authors performed an analysis of over 60 days of operational TDRS Unused Time reports for three TDRS nodes, comprising more than 3,700 blocking periods, to estimate the probabilistic factors that influence access wait time for the event-driven method[5]. Each TDRS node selected for this analysis is stationed in a unique service region to ensure global coverage. These nodes constitute the most suitable subset of the larger TDRS constellation for transient science applications and are currently used by Swift. Each TDRS node has the capacity to support several simultaneous service instances for multiple-access return services. In the dataset analyzed, the TDRS nodes always had unscheduled capacity for least one multiple-access return

[5] An analysis of the available period distribution was also performed to ensure that when a resource is (or becomes) available, it is available for the required duration of the requested service period. Swift currently schedules a four minute service period to send and verify receipt of follow-up commands. The probability of an available period being equal to or greater than four minutes was found to be 98.7%. The results reported in this chapter ignore the small effects of the percentage of available periods that are not long enough to satisfy the minimum user service period duration. Note, because the service period duration is a user-specified parameter in the SAR, SN systems will never allocate available periods that are too small to complete the user's operation. As a result, available but unusable periods manifest as slightly longer (i.e., by the duration of the available but unusable period) blocking periods than would be manifest in the TDRS Unused Time dataset. Improving user flight and ground software synchronization times and user radio signal acquisition performance would minimize the likelihood and impact of this occurrence.

service instance. Blocking occurred for multiple-access forward services only, as each TDRS node can presently support a single service instance of this type at a time. As a result, blind commanding or using the continuously available DAS for verification that the sent command was received does not provide any improvement in service access wait time. The probabilistic factors associated with SN blocking periods were derived from the TDRS Unused Time dataset and are as follows: the mean constellation probability of blocking (*P[Blocking]*), the expected duration of a blocking period (*E[Blocking]*), and the expected wait time contribution due to blocking periods (*E[Blocking]/2*).

The analysis of TDRS blocking periods revealed the presence of a number of long duration blocking periods greater than 100 minutes. These long duration blocking periods are associated with SN operational scenarios that do not involve LEO user services, such as SN maintenance and sustainment, pre-launch user compatibility testing and high-altitude balloon campaigns. TDRS nodes have substantial coverage overlap allowing SN planners to minimize the impact of long duration blocking periods on the overall SN service availability. Additionally, LEO missions such as Swift will transit the coverage area of all TDRS nodes in the constellation over the course of each orbit, providing at least three opportunities to obtain service access approximately every 90 minutes. As a result of these factors, the long duration blocking periods increase the expected (mean) access wait time contribution due to blocking periods beyond what is likely to be experienced by LEO users. The median blocking period and associated access wait time contributions were calculated to quantify the degree to which the mean access timeliness metrics are impacted by the long duration blocking periods. The median wait time due to blocking was found to be 28 seconds less than the mean wait time. A summary of the probabilistic factors

associated with SN blocking periods, including factors calculated with the median blocking duration, are provided in Table 7 below.

Table 7: Summary of probabilistic factors associated with SN blocking periods.

Expected Duration of Blocking *E[Blocking]* (H:MM:SS)	**Expected Wait Time Added Due to Blocking *E[Blocking]/2* (H:MM:SS)**	**Median Duration of Blocking (H:MM:SS)**	**Median Wait Time Added Due to Blocking (H:MM:SS)**	**Probability of SN Resource Blocking *P[Blocking]* (Mean, 3-node Constellation)**	**Probability of SN Resource Availability *P[Available]* (Mean, 3-node Constellation)**
0:05:55	0:02:58	0:05:00	0:02:30	16%	84%

If the SN resource is blocked, the wait time due to blocking is added to the additional wait time contributors identified in Eq. (7). If the requested SN resource is available, the minimum service access wait time is achievable, and comprised of the factors identified in Eq. (4). A summary of the access wait time outcomes for the three SAR build options are provided in Table 8 below.

Table 8: Summary of SN event-driven method service access wait time outcomes for three SAR build options.

Schedule Add Request Build Option	**Minimum Access Wait Time, $Access_{Min}$ (H:MM:SS)**	**Expected Access Wait Time If Blocking Occurs, $E[Access_{Blocking}]$ (H:MM:SS)**	**Median Access Wait Time If Blocking Occurs (H:MM:SS)**
SNAS Graphical User Interface	0:10:40	0:14:08	0:13:40
Prototype Schedule Add Request	0:10:15	0:13:43	0:13:15
External Processing System (Fully Automated)	0:10:05	0:13:33	0:13:05

Finally, the total expected access wait time for the event-driven method is calculated as a weighted average sum of the access wait time outcomes, accounting for the probability that SN resources are either available or blocked when service access is requested, as described in Eq.

8. The total expected access wait time results for the event-driven method are summarized in Table 9 below.

Table 9: Total expected access wait time for mean and median blocking contributions and each SAR build option.

Schedule Add Request Build Option	**Total Access Wait Time for the Event-Driven Method – Using Expected Blocking Wait Time, $E[Access_{EventDriven}]$ (H:MM:SS)**	**Total Access Wait Time for the Event Driven Method – Using Median Blocking Wait Time (H:MM:SS)**
SNAS Graphical User Interface	0:11:13	0:11:09
Prototype Schedule Add Request	0:10:48	0:10:44
External Processing System (Fully Automated)	0:10:38	0:10:34

Discussion of Results

The SN provides global communications coverage for LEO users, in contrast to the intermittent coverage provided by direct-to-earth ground networks. Transient science missions have relied on manual intervention by SN operators to obtain more timely service access than is possible from fixed ground network service periods. The results of this research provide an evaluation of two complimentary methods for improving service access timeliness outcomes using SN. Table 10 summarizes and compares the state-of-the-practice, gap-filling and event-driven service access methods.

Table 10: Summary of SN timely service access methods.

	State-of-the-Practice	**Gap-filling Method**	**Event-driven Method**
Method Implementation	Relies on manual intervention by user and network operators	Relies on the frequency of pre-determined service periods from ground and SN provider networks	Relies on novel SN procedures to quickly access available services
Implementation Timeframe	During real-time operations	During the nominal schedule pre-planning phase	During real-time operations
Degree of Automation	Manual-only	Partial or full automation	Partial or full automation
Dependability	Best-effort due to blocking, staffing and other constraints	Guaranteed based on mission priority	Best-effort due to blocking and first-come-first-served scheduling policy
Timeliness Outcomes	At least 20 min. following service access request	Depends on the magnitude and number of fixed schedule service gaps	~10 min. wait time achievable 84% of the time; expected wait time < 14 min. if blocking occurs
Application Suitability	Spacecraft emergencies; not suitable for nominal transient science operations	Suitable for short-term use to validate new observation workflows or measurements; suitable for long-term use to mitigate timeliness impacts of the longest ground gaps	Suitable for long-term transient observation campaigns or for rarely occurring transient observables

Traditional space communications service planning and operations processes introduce systemic inefficiencies when users have emergent time-sensitive operational needs. These inefficiencies are amplified in transient science applications, where multiple users must interact and collaborate in dynamic workflows. The systemic inefficiencies can be attributed to the inability of current processes to adequately accommodate random user traffic needs, the inability

of users to specify goal-oriented service access parameters to achieve desired outcomes and the lack of shared or coordinated service planning and information among provider networks.

The gap-filling method pathfinder experiment illustrates many of these inefficiencies. Swift operators were required to plan and submit two separate batches of service access requests to the ground network and SN for operations occurring 12 to 19 days in the future. While Swift required the ground network service instances to perform nominal Swift operations, the gap-filling SN service instances were scheduled to reduce the expected service access wait time just in case an unplanned follow-up operation was needed. Each SN service access request included parameters specifying a specific TDRS node at a specific time, setting up a binary resource allocation choice for SN schedulers: either a higher-priority user requested the same resource at the same time or it is available. User adoption of goal-oriented service access requests incorporating flexibility in the specific service delivery node, service start time, or service duration would reduce the impact of SN resource blocking and improve the expected wait time for service access. The SN offers basic flexible scheduling capabilities, but these features are not widely known or used at the present time. Widespread adoption of goal-oriented service access requests would allow network service providers more discretion to efficiently allocate service delivery tasks to network resources across partnered commercial or other federated service providers. Future research is planned to understand the lack of user adoption of SN flexible scheduling features, which may reveal a presently unknown barrier to the emergence of a robust network provider ecosystem for LEO users. Finally, the two duplicative service access requests submitted by Swift operators for overlapping ground and SN service periods illustrates user and

network planning burden and resource allocation inefficiencies that could be addressed by improved coordination of service planning information among provider networks.

Results from the event-driven pathfinder experiment demonstrate that the 10-minute SN minimum lead time to service period start, as specified in SN documentation, dominates the expected service access wait time contributions from all other sources. Testing performed on the ANCC revealed that the minimum lead time factor is comprised of a fixed 6-minute component and a variable, but small delay associated with SAR processing. As a result, it may be possible to reduce the minimum lead time by up to 40% without modification to current SN systems. Future work to operationally characterize the variable processing delay and to validate the fixed component of the SN minimum lead time could result in improved access timeliness beyond the results reported in this chapter. Furthermore, the fixed 6-minute component of the minimum lead time appears to be set by the worst-case inter-user pointing time for the TDRS mechanically steered antenna used for single-access services. The event-driven method relies on the electronically steered multiple-access TDRS phased array resources, which have a specified and experimentally verified inter-user setup time of 30 seconds. This suggests that the minimum service lead time may be reduced by up to 95% if the SN systems could filter time-sensitive event-driven multiple-access service requests from requests for the single-access services.

The access wait time outcomes of the event-driven method provide substantial improvements in effectiveness and efficiency when compared to the reported 20-minute minimum service access wait time achievable with the current manual intervention method. The minimum event-driven access wait time outcome is a 47-50% improvement to the manual intervention method depending on the SAR build option selected by the user. The largest

expected service access wait time outcome of a little over 14 minutes, which occurs if blocking is present and SARs are built using the graphical interface, is also an improvement over the manual intervention method. Both event-driven method timeliness outcomes satisfy Swift's gravitational wave follow-up scenario threshold of 30 minutes. Additionally, the event-driven method reduces or eliminates real-time operator burden, depending on the user selected SAR build option. However, an increase in the number of transient science observatories conducting simultaneous follow-up operations will increase the probability of resource blocking, leading to longer user access wait time outcomes. Future work is needed to explore the relationship between the traffic demands of larger-scale multi-observatory transient science workflow concepts and SN service delivery capacity. A summary of the SN access wait time contributors identified by this research and a roadmap for possible improvements is provided in Fig. 8.

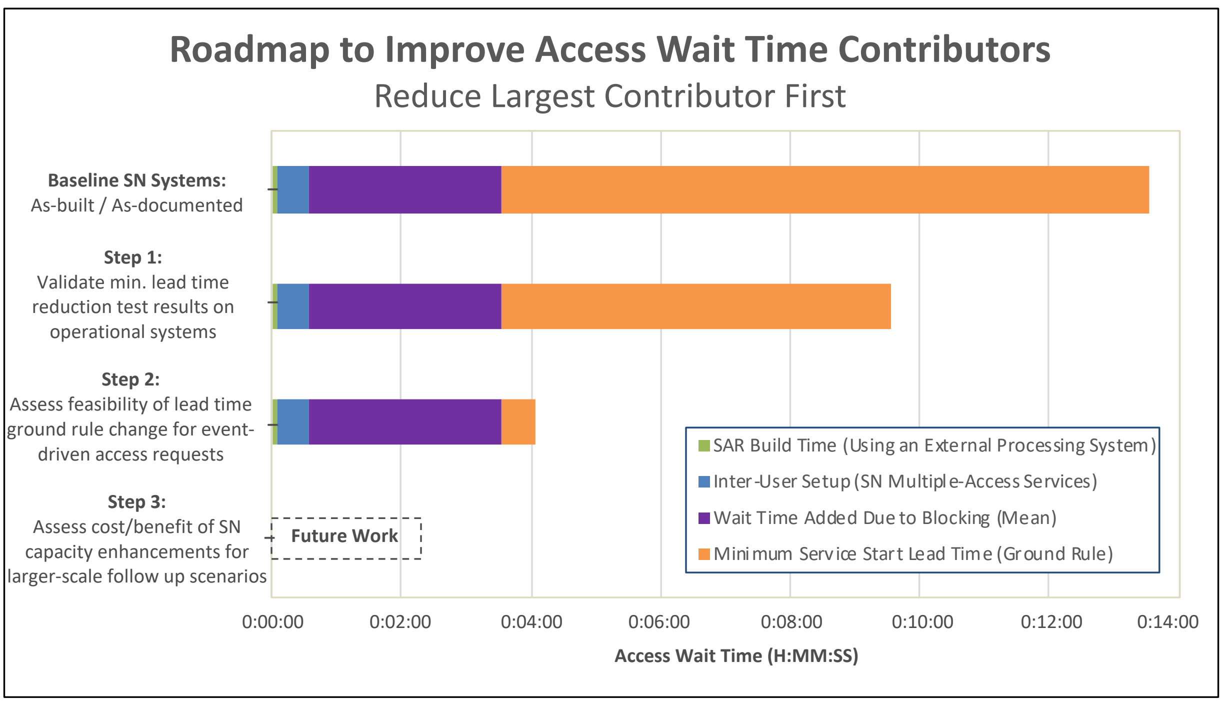


Fig. 8: Roadmap to improve SN access wait time contributors for the event-driven method.

Conclusion

The two service access methods developed by this research can be used in complimentary ways to mitigate the systemic inefficiencies of traditional service planning processes for time-sensitive mission operations. Transient science missions can apply the results from this research to automate the analysis of weekly ground network service gaps and identify the number and timing of SN gap-filling service periods that would achieve the desired average service access wait time. Then, the gap-filling batch of SN service access requests could automatically be built and submitted using an External Processing System. This process results in a guaranteed expected service access wait time outcome, which must be balanced with the additional loading imposed on SN resources through negotiations among mission and network stakeholders. For operational scenarios where the gap-filling method is infeasible, the event-driven method can be invoked during the active period to provide users with best-effort access to services using the available (unreserved) SN capacity. Test results indicate that the baseline SN systems can provide event-driven service access in as little as six minutes, with the possibility of reducing access wait times to under one minute if modifications to the processing priority of SN multiple-access service access requests are made.

In summary, the results from this research indicate that the as-built SN systems have adequate capabilities and capacity to support infusion of the gap-filling and event-driven methods to current and future users with time-sensitive mission operational needs. These methods are consistent with and traceable to NASA's overarching network and mission architecture. Process and technical options to substantially improve user service access timeliness outcomes beyond the results reported in this chapter were identified. Future work is

planned to better understand the impact to access timeliness outcomes caused by the dynamic traffic loading associated with multiple interacting space-based observatories. In addition, future opportunities to conduct pathfinder experiments incorporating commercial service providers are envisioned to meet the increasing need for network services enabling distributed time-sensitive space mission operations.

Chapter 4

Chapter 4 addresses Research Question #2:

How can the publish-subscribe messaging design pattern be applied to space data networks and its impact be evaluated for transient science space systems?

Introduction

The study of transient scientific phenomena using diverse ground and space-based observatories presents new opportunities for scientific breakthroughs [1] [2]. Many of the most relevant transient scientific events emerge randomly and require prompt observations to obtain data of the highest scientific value [3] [4] [5]. The activities of multiple observatories may be structured into purposefully designed transient science observation scenarios whose successful execution depends on timely data flows across terrestrial and space communications networks [6].

Multipoint messaging applications allow dissemination of data from one source to multiple destinations in a communications network [18]. Communications network users register their preferences about message topics and data contents using a publish-subscribe relational model [18]. Space-based observatories presently provide timely notifications of transient events to ground-based mission operators using the continuously available space-to-ground data transport services of the NASA Near Space Network's Tracking and Data Relay Satellite (TDRS) space relay communications network [32]. Notifications of transient event observations are subsequently disseminated by multipoint data flows to the scientific community using transient

science messaging applications such as the Gamma-ray burst Coordinates Network (GCN) over terrestrial Internet Protocol (IP)-based communications networks [41] [17]. Some ground-based observatories have implemented machine-to-machine interfaces to the GCN to automate timely follow-up observations.

The TDRS space data network and the GCN transient science messaging application allow space-based observatories to initiate multi-observatory transient science observation scenarios. However, space-based observatories cannot presently be connected to transient science messaging applications to perform timely follow-up observations due to limitations in ground-to-space communications network infrastructure, protocols and autonomy [42] [43] [35] [44]. In addition, current design and evaluation methods do not account for important considerations raised by multi-observatory transient science observation scenarios, such as the impact of correlated user demand for space data network services and the amplification of data traffic induced by the scenario, among others. Science mission and network planners lack a systematic framework to formulate and evaluate solution options for implementing multipoint space data flows in transient science observation scenarios.

The three goals of this research are: (1) to define two functional reference architectures for implementing timely multipoint ground-to-space data flows, (2) to provide a multipoint space data flow design and evaluation method for transient science observation scenarios and (3) to apply the functional reference architectures and multipoint space data flow design and evaluation method in a notional transient science space system design study.

Concept and System Model Development

Section 1 (this section) defines two functional reference architectures for implementing timely multipoint ground-to-space data flows. The *command pipeline* and *notification pipeline* provide mission and network planners with alternative solution templates for the design of transient science observation scenarios involving multiple space-based observatories.

Concept Definition

A transient science space system is comprised of at least two observatories – with at least one being space-based – whose operations are connected by time-sensitive data flows. Mission and network planners purposefully design a transient science space system by creating a model that defines one or more ways in which the functions and information flows among the system elements can fulfill its scientific objectives [45]. Functional reference architectures encode a particular allocation of system functions and information flows that can be used as a solution template for common design problems [24]. Alternative functional reference architectures facilitate design exploration and evaluation analyses because they have differing dependencies, risks and other characteristics used by planners to judge the suitability of a design solution.

The *command pipeline* and *notification pipeline* functional reference architectures provide two alternative approaches to achieve the timely dissemination of data from the ground to multiple space-based observatories. The reference architectures describe a functional interface to a terrestrial multipoint science messaging application within an end-to-end multi-observatory transient science observation scenario. Each reference architecture can be implemented to automate service management data flows, which allow users to signal their timely data transport needs to the space data network, and service execution data flows, which

transport user data from a ground-based source to space-based destinations. The command pipeline and notification pipeline reference architectures are represented using the conventions of Systems Modeling Language (SysML). Data flow coordination and sequencing control considerations for each reference architecture are also discussed.

Command Pipeline Functional Reference Architecture

The command pipeline serves as a functional interface to the terrestrial transient science messaging application. It is implemented at the mission operations centers of space-based observatories involved in a transient science observation scenario. It is subscribed to messages from the transient science messaging application and automates follow-up observation decisions according to the users' preferences. The command pipeline automates traditional mission control functions, including requesting access to the space data network, building observatory commands in accordance with mission security policies and sending the command data to the to the space-based observatory at the reserved space data network access period.

End-to-End Operational View

An end-to-end operational view of the data flows for a transient science observation scenario involving command pipelines at several space mission operations centers (MOCs) is provided in Fig. 9.

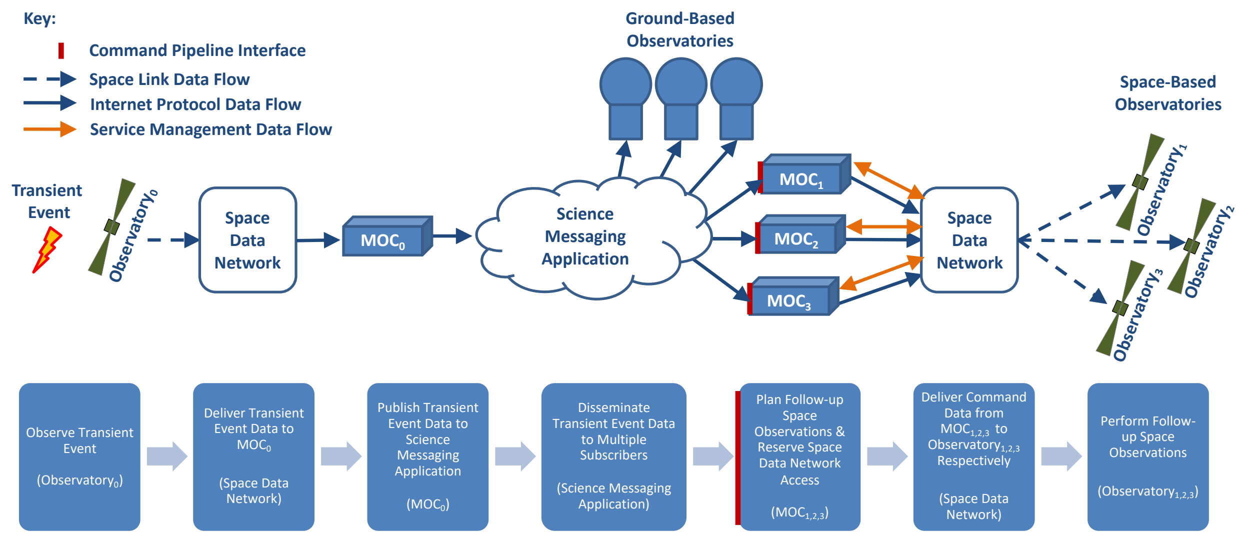


Fig. 9: Operational view of a transient science scenario using command pipelines to implement follow-up observations.

The operational flow in Fig. 9 begins when a transient scientific event is detected by the initial space-based observatory ($Observatory_0$) in the transient science observation scenario. Information about the transient event from $Observatory_0$ is sent to the ground via a timely space-to-ground data transport service, such as the TDRS Demand Access Service, or at a pre-reserved space data network access time. The space data network sends information from $Observatory_0$ to its mission operations center (MOC_0) using an interface to an IP-based terrestrial network. At MOC_0 (or an associated science operations center), the information from $Observatory_0$ is processed into a data product. This data product will be distributed as a message to multiple destinations according to pre-defined preferences from the scientific community and operators of other observatories. Preferences could include the desired notification topics, message data formats, and timeliness (e.g. as soon as possible, daily summary etc.). A science messaging application applies these preferences and uses an IP-based network to implement multipoint data flows to many ground and space-based observatory control centers simultaneously. To achieve operational efficiency and timeliness, some ground-based observatories have

implemented interfaces and control algorithms to allow the data product from the science messaging application to initiate autonomous follow-up observations.

Continuing to follow Fig. 9, the command pipeline serves as the functional interface to the science messaging application. A command pipeline implemented at each mission operations center (MOC_1, MOC_2, MOC_3) performs planning and specification of the follow-up operations for its associated observatory. Specifically, the command pipeline applies the decision rules and constraints, builds or selects appropriate software commands encoding the desired follow-up operations, and obtains timely access to the space data network link to send commands to its space-based observatory. The space-based observatory can perform its specified follow-up operations once it receives command data from the command pipeline implemented at its MOC.

An amplification of the demand for space data network services can occur when the follow-up operations scenario involves multiple space-based observatories. The timeliness with which a space data network can deliver commands to all observatories is a performance metric useful for assessing the feasibility of multi-observatory transient science scenarios and in evaluating the overall suitability of alternative design solutions.

In summary, the command pipeline serves as a functional interface to the science messaging application within an end-to-end transient science observation scenario. A command pipeline is implemented at each mission operations center involved in a transient science observation scenario. It is responsible for the service management data flows associated with obtaining timely access to space data network links and for building, securing and sending the command data to the space data network at the reserved space data network access time. The

command pipeline interface (red), ground-to-space service management (orange) and service execution (blue) data flows are illustrated in Fig. 9.

The end-to-end operations represented in Fig. 9 are intended to serve as a template for the composition of more complex transient science observation scenarios. For example, Observatory_1 which serves as a data destination in the above scenario may become the data source (Observatory_0) in a multi-tiered operations scenario. More complex scenarios may be designed that apply sequencing rules among observatories or that involve conditionally branching pathways that would be activated based on timely indicators in the science message data stream. The design of such complex scenarios can expand upon the simple example provided here.

System Structure

The structural model of a transient science space system developed in Chapter 3 and in Roberts et al. has been extended to illustrate the relationship of the command pipeline within the system hierarchy [46]. Specifically, the command pipeline is represented as a component of the mission control element within the user segment of a transient science space system. The science community external actor represents the source observatory and science messaging application depicted in the end-to-end operational view in Fig. 9. The user segment represents one or more users of the space data network. The network service provider segment represents a single space data network for simplicity. However, NASA is increasingly investing in technologies that would allow for interoperability with many emerging commercial space data network service providers. Future work is planned to address the implications of commercial

space relay and direct-to-Earth space data network service providers for transient science space systems. The command pipeline is illustrated within the system hierarchy in Fig. 10.

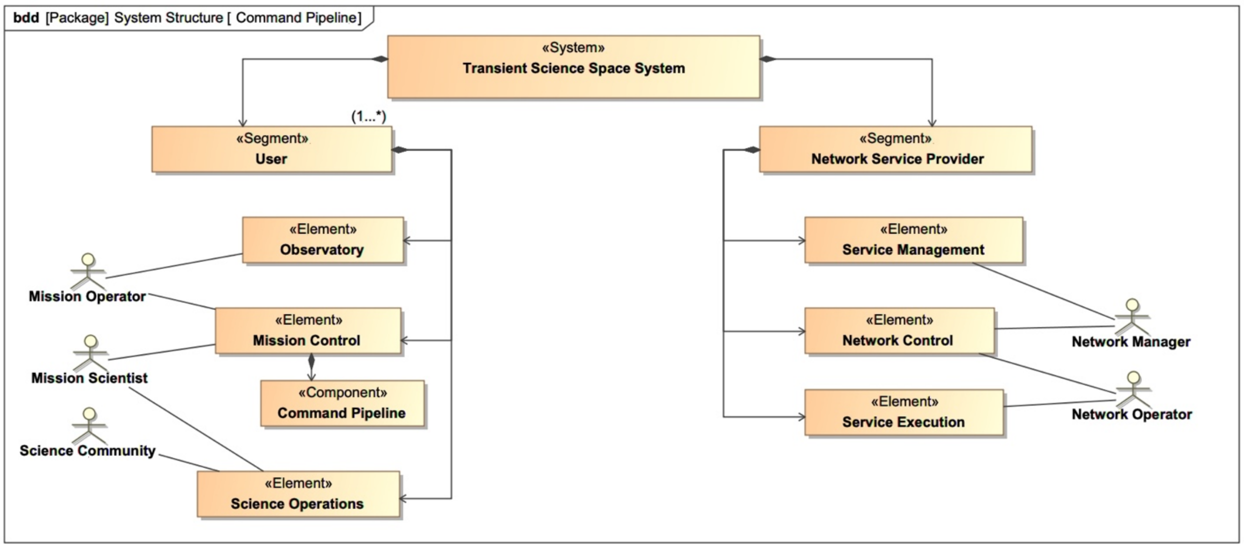


Fig. 10: Structural model of a command pipeline within the overall transient science space system hierarchy.

Activity Flow and System Allocation

The command pipeline functional architecture automates mission control functions that have traditionally been performed by manual procedures to improve the efficiency and timeliness of follow-up observations of transient events. The command pipeline activity flow and allocation within system structure are illustrated in Fig. 11.[6]

[6] The service management activity flows acknowledging receipt of a user's service access request and confirming or rejecting the requested access times are not shown for clarity.

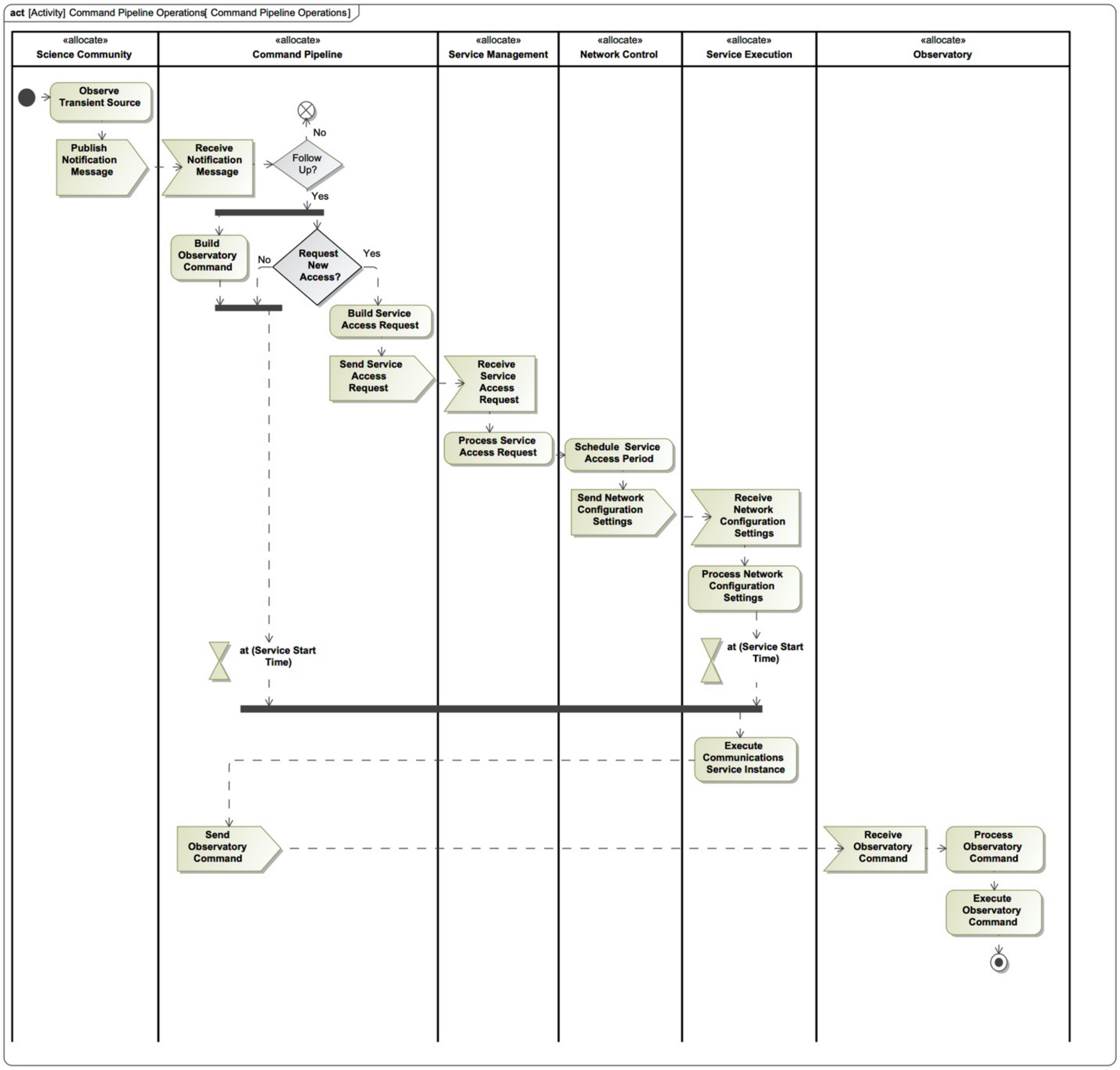


Fig. 11: Activity flow of a command pipeline allocated to elements in the system structure.

As depicted in Fig. 11, the activity flow originates with the scientific community external actor. The science community external actor represents the source observatory and science messaging application depicted in the end-to-end operational view in Fig. 9. The scientific data is received by the command pipeline component within the mission control element of the user segment. The command pipeline implements a follow-up decision algorithm that evaluates the

relative scientific priority, operational feasibility and engineering constraints associated with follow-up observations of the transient source described in the data received from the science messaging application. In the affirmative case, the activity flow splits into two parallel paths. One path leads to the software build or selection of the appropriate commands specifying the sequence of follow-up operations to be performed by the observatory. Generally, commands are unique to each observatory and have low data volume, ranging from as little as a few bits up to a few kilobits. Once the command is built or selected, it is stored until the service execution start time of the space data link. The other path leads to a second decision node, which evaluates whether a new service access request should be submitted to the space data network to improve the timeliness of the command delivery to the observatory. In cases where the decision to perform follow-up operations depends on the achievable timeliness that the command can be delivered to the observatory, it may be necessary to incorporate feedback between the service access and follow-up decision nodes. In cases where users have pre-scheduled space data network access, the second decision node evaluates whether the wait time until the next pre-scheduled service is acceptable. In these cases, it is not necessary to initiate a new space data network service access request and the activity flow merges at completion of the command build activity. If more timely access to space data network services is necessary, the command pipeline initiates a service access request, which is sent to the network service provider's service management element.

When designing transient science observation scenarios with command pipelines, it is important to consider how the intended cross-observatory coordination behaviors will be implemented. This may involve the need for additional decision algorithms or activities across

user-to-user or user-to-network service provider interfaces. For example, many space data network services are reserved by other users weeks in advance for pre-scheduled operations. As a result, timely service access may be subject to latencies induced by the blocking of specific resources. To improve the probability of successfully reserving timely space data network services, the service access request may include flexibility or goal-oriented parameters, such as windows of acceptable service start times spanning multiple service providing resources. Space data networks may also publish a registry of unused times, which can be ingested and used by the command pipeline to inform the parameter specification for a new service access request (e.g. only request access for available resources). For network service providers that employ a first-come-first-served scheduling policy for unused resources, such as the TDRS space data network, some coordination of the service access request specification parameters may be needed across user command pipelines to ensure the appropriate priority and sequencing of service access among the observatories in the scenario. This coordination is needed to avoid service access priority inversions – a condition where a higher priority user is blocked by a lower priority user simply because the lower priority user's service access request was received first. Network service providers with sophisticated scheduling capabilities could implement the sequencing constraints and access priorities necessary to implement a multi-observatory scenario.

Once the service access request is received by the network service provider, a series of activities involving the reservation of resources, configuration of resources and transport of space data are performed by the service management, network control and service execution elements, respectively.

At the service access start time, the command is sent from the command pipeline and received by the observatory. The observatory subsequently processes the command and executes the follow-up operations specified by the commands.

Summary

The command pipeline functional reference architecture maintains the traditional system partitioning and functional allocation between the user and network service provider segments. The command pipeline activity flow describes the automation of user mission control and network service management functions to achieve improved operational efficiency and timeliness. To implement a command pipeline, the network service provider's service management element must have the capability to accept short lead-time user service access requests and have adequate capacity to execute services on timescales appropriate for the transient science scenario. The TDRS space data network satisfies these preconditions [46]. The network service provider's scheduling capabilities should be considered for scenarios involving multiple space-based observatories. Coordination among user command pipelines may be necessary to implement any sequencing or access priorities required by the scenario when the network service provider uses a first-come-first-served scheduling policy.

> In collaboration with this research, the Swift mission implemented the first known command pipeline in January 2020 [17]. As a result, Swift has increased its expected detection rate of electromagnetic counterparts to ground-observed gravitational waves by greater than 400%.

Notification Pipeline Functional Reference Architecture

Traditional space mission analysis and design practices identify two classes of communications data traffic: (1) mission data (i.e. data from the instrument payload) and (2) engineering data (i.e. spacecraft tracking, telemetry and command data) [20]. This research highlights *multipoint event notification data* as a third increasingly important class of space data traffic. Notification data are messages comprised of mission, engineering, environmental and other operational event status and state change information. Notification messages are disseminated by multipoint data flows connecting users and network service provider nodes. Notification messages are used to synchronize and orchestrate activities across space-based observatories and with the space data network. The timeliness requirements for notification data flows are driven by the higher-level synchronization and orchestration needs of a particular application or operational scenario. Notification data are implemented as standardized messages to mitigate the potential demand amplification and traffic amplification impacts of multipoint data flows. The asynchronous messaging standard published by the Consultative Committee on Space Data Standards (CCSDS) has been identified in the literature as one suitable approach to implement standard notification messages for space systems [47]. The security architecture for users of notification data is consistent with the science and multi-organizational spacecraft mission profiles described in the CCSDS Security Architecture for Space Data Systems Magenta Book [48]. A more detailed discussion of information security functions and implementation considerations applicable to notification pipelines can be found in Sanchez Net et. al. [49]. Although the notification pipeline reference architecture is suitable for a wide variety of

applications, this section focuses on its application to multi-observatory transient science observation scenarios.

The notification pipeline serves as a functional interface to the terrestrial transient science messaging application and provides a subscription-based content delivery service to space-based observatories. It is implemented by the space data network service provider as an extension to traditional network service management functions. The notification pipeline is subscribed to messages from the transient science messaging application. It applies pre-defined user preferences about information received from the terrestrial science messaging application and orchestrates multipoint space data flows transporting the notification message to space-based observatories. The notification pipeline converts and applies space data standards and protocols to data received from the terrestrial science messaging application to mitigate the impacts of correlated service demand and traffic amplification in multi-observatory transient science scenarios. The notification pipeline reference architecture requires reallocation of some functions traditionally performed by the user mission control element to the observatory. In particular, the observatory must have sufficient autonomy to respond with appropriate follow-up observation behaviors based on the receipt of standardized common notification messages, in contrast to the unique observatory-specific commands used in command pipelines.[7] As compared to the command pipeline reference architecture, the notification pipeline may simplify user access, improve data timeliness and offer better scalability to large numbers of subscribed

[7]Observatory autonomy is also needed for information security, fault-detection, and other supporting functions to enable appropriate follow-up observation behaviors.

observatories depending on the design details of the multipoint data flows and the capabilities of the space data network.

End-to-End Operational View

An end-to-end operational view of a multi-observatory transient science observation scenario involving a notification pipeline is provided in Fig. 12.

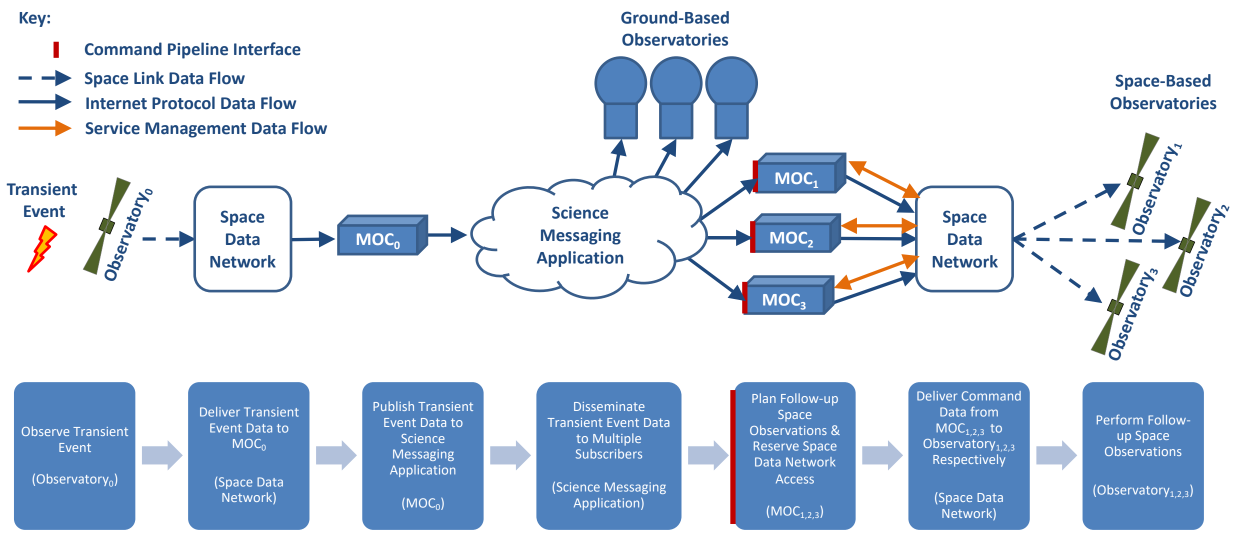


Fig. 12: Operational view of a transient science scenario using a notification pipeline to implement follow-up observations.

Within the end-to-end transient science observation scenario depicted in Fig. 12, the notification pipeline serves as a functional interface between the science messaging application and the space data network. Users of the notification pipeline (MOC_1, MOC_2, MOC_3) pre-register their service management preferences with the space data network. These preferences could include subscription attributes such as the desired topics or source observatories, observable thresholds indicated in the science message data stream, and delivery timeliness or sequencing requirements among multiple space-based observatories. This information is used by notification

pipeline algorithms to reserve and execute space data network services in accordance with the users' preferences and the multi-observatory transient science observation scenario objectives.

Compared to the multipoint data flows of command pipelines, notification pipelines are intended to diminish the impacts of correlated service demand and traffic amplification by disseminating common standardized data products for use by multiple space-based observatories. Although the data volume of common standardized notification messages may be comparable to that of a unique command, the notification pipeline functional architecture does not preclude the dissemination of high-fidelity science data products directly among collaborating space-based observatories. However, such scenarios would demand greater degrees of available network bandwidth and observatory onboard processing. The usage of command pipelines or notification pipelines in multi-observatory transient science observation scenarios has an influence on several system suitability metrics that will be discussed and quantitatively explored in subsequent sections of this chapter. To enforce common data standards across multiple space-based observatories, the notification pipeline performs any necessary processing to convert the terrestrial science messaging application data products into the appropriate space data standards and protocols before disseminating the notifications to the subscribed observatories. More details about candidate space data standards and protocol stack recommendations for implementation of a notification pipeline can be found in references [47] and [50].

System Structure

The structural model of a transient science space system developed in Roberts et al. has been extended to illustrate the relationship of the notification pipeline within the system

hierarchy [46]. Specifically, the notification pipeline is represented as a component of the service management element within the network service provider segment of a transient science space system. The science community external actor represents the source observatory and science messaging application depicted in the end-to-end operational view in Fig. 12. The user segment represents one or more users of the space data network subscribed to a notification pipeline. As with the command pipeline, a single network service provider is considered for simplicity. The notification pipeline is illustrated within the system hierarchy in Fig. 13.

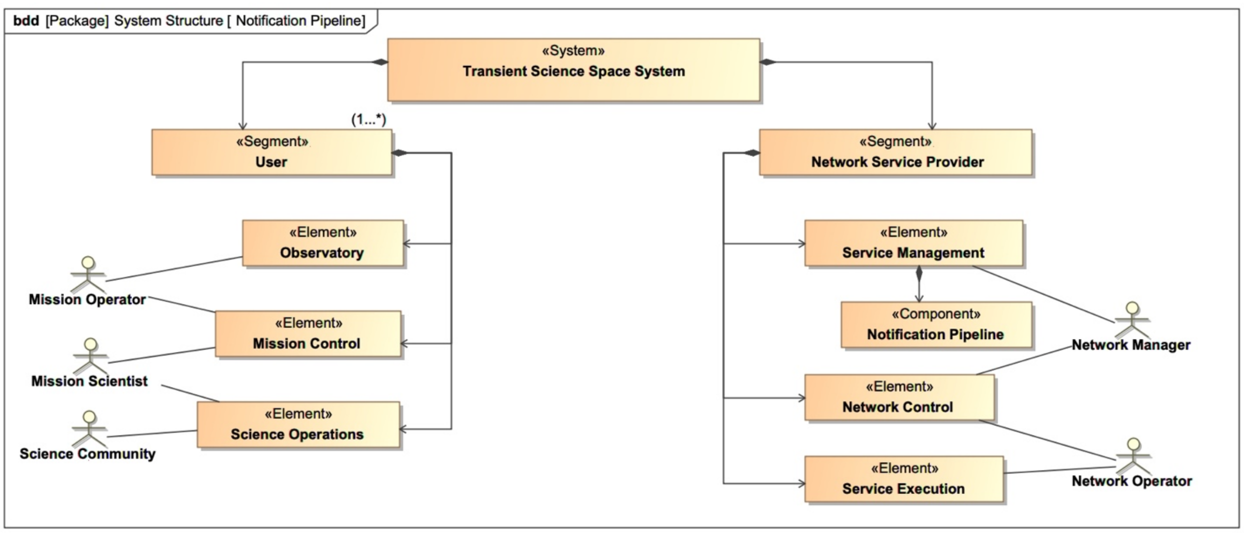


Fig. 13: Structural model of a notification pipeline within the overall transient science space system hierarchy.

Activity Flow and System Allocation

The notification pipeline functional architecture enables space-based observatories to perform timely and automated follow-up observations by interfacing terrestrial transient science messaging applications with space data networks. The notification pipeline activity flow and allocation within the system structure are illustrated in Fig. 14.

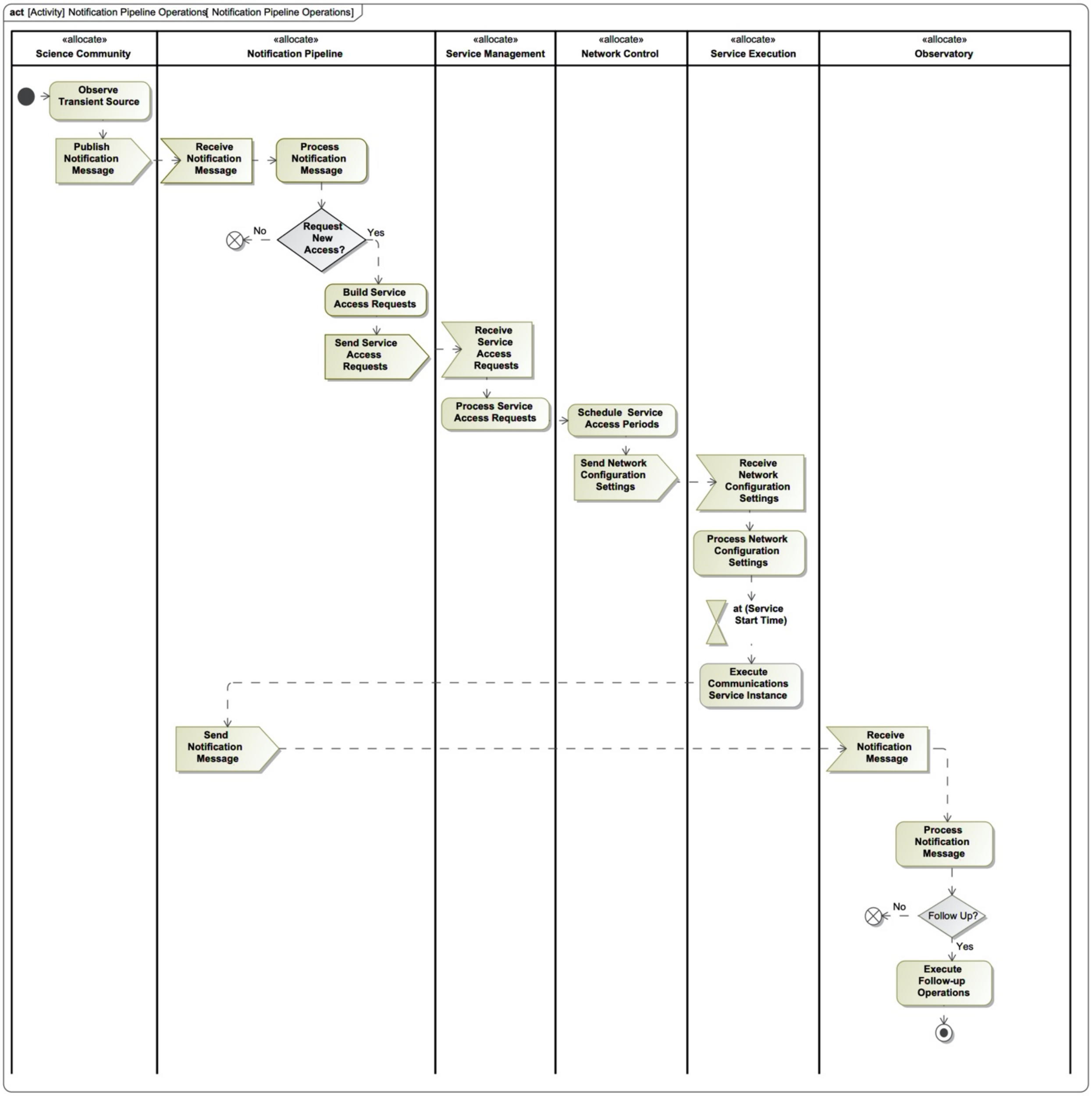


Fig. 14: Activity flow of a notification pipeline allocated to elements within the system structure.

As depicted in Fig. 14, the activity flow originates with the scientific community external actor. The science community external actor represents the source observatory and science messaging application depicted in the end-to-end operational view in Fig. 12. The scientific data is received by the notification pipeline component within the service management element of

the network service provider segment. The notification pipeline processes and converts the terrestrial science message application data to comply with space data messaging application standards and protocols. The notification pipeline implements a decision algorithm to assess whether the scientific data message received form the terrestrial science messaging application satisfies the pre-determined user preferences and constraints for autonomous dissemination to one or more space-based observatories. If so, the notification pipeline requests space data network service access as a proxy for each subscribed users' mission control element. The notification pipeline may also implement the scheduling capabilities necessary to achieve any sequencing order or access priorities among the subscribed space-based observatories involved in a complex multi-observatory transient science observation scenario.

The service access requests generated by the notification pipeline for each subscribed user are submitted to the traditional service management functions of the network service provider. A series of activities follows involving the reservation of services, configuration of resources, and data transport functions, performed by the traditional service management, network control and service execution elements, respectively.

At the service access start time, the common notification data is sent from the notification pipeline and received by one or more space-based observatories. Each observatory subsequently processes the standardized space notification message data and evaluates whether and how to perform follow-up operations to carry out its role in the transient science observation scenario. If so, the observatory executes the appropriate follow-up operations.

Summary

The notification pipeline functional architecture enables space-based observatories to perform timely and automated follow-up observations by interfacing terrestrial transient science messaging applications with space data networks. It requires changes to the traditional system partitioning and functional allocation between the user and network service provider segments. The notification pipeline expands traditional network service management functions by applying pre-specified user preferences about the dissemination of terrestrial messaging application content, converting terrestrial messaging data to appropriate space messaging and security data standards and protocols and reserving access to space data links for the transport of notification messages on behalf of the user. In addition, space-based observatories must have sufficient autonomy to respond with appropriate follow-up observation behaviors based on the receipt of standardized notification messages, in contrast to the unique observatory-specific commands used in command pipelines. As compared to the command pipeline reference architecture, the notification pipeline may simplify user access, improve data timeliness and offer better scalability to large numbers of subscribed observatories, especially if the multipoint data flows are implemented by physical space broadcast links. The TDRS space data network has experimentally demonstrated a suitable implementation of a physical space broadcast link [36].

Presently, there are no instances of multipoint space messaging notification pipelines in operation. However, the notification pipeline functional reference architecture is compatible with TDRS and other space data networks, provided they have the capability to accept short lead-time user service access requests and have adequate capacity to execute services on timescales appropriate for the user's operational scenario. Furthermore, the Swift observatory is presently able to perform similar onboard autonomy tasks called for by the notification pipeline functional reference architecture. Specifically, Swift's follow-up decision algorithm can evaluate the relative priority of interrupting its pre-planned observations with the narrow field-of-view X-ray and optical instruments when its wide field-of-view Gamma-ray burst instrument detects a transient source. Within several seconds of detection, the Swift observatory may slew to the celestial coordinates of the detected Gamma-ray burst and begin unplanned observations at the X-ray and optical wavelengths. In principle, it should not matter to the Swift follow-up decision algorithm whether the Gamma-ray burst was detected by an instrument co-manifested on the same spacecraft bus or whether the notification comes from another observatory. In short, there are mature technical and operational foundations in place for the implementation of notification pipelines.

Multipoint Space Data Flow Design and Evaluation Method Development

Command and notification pipelines provide a means to use space data networks to achieve the automated multipoint data flow patterns found in terrestrial science messaging applications. However, quantifying the effects of command and notification pipelines in transient science systems pose several new challenges to current space mission design and analysis

practices. For example, coordinated user behaviors in a transient science observation scenario impact both the demand for access and traffic loading of space data networks. Transient science observation scenarios may impose requirements for access precedence and data delivery sequencing among users that are beyond the capabilities of current space data network scheduling capabilities. Finally, utilization of space data network services is impacted by the frequency with which the transient science observation scenario is executed. Section 2 (this section) provides a mathematical framework and transient science observation scenario modeling guidelines to address each of these challenges.

Theoretical Framework

Queueing theory is concerned with the mathematical study of the delays of waiting in line [39]. Queueing models have been applied to inform system design and analysis in diverse engineering, manufacturing and logistics contexts. Queueing models provide a quantitative framework for relating queue arrival processes and servicing processes to system timeliness, capacity and throughput outcomes [39] [51]. As such, queueing theory provides a well-established basis for modeling, analysis and evaluation of data flows in transient science space systems.

Queue Arrival Process Modeling Guidelines

In defining the inputs to this process, we first consider the space data network service request arrival processes within an end-to-end transient science observation scenario. Transient scientific events occur randomly in space and time. An initial observation from a space or ground-based observatory initiates the transport of science data products within terrestrial transient science messaging application, as depicted in the command and notification pipeline operational

views, Fig. 9 and Fig. 12 respectively. Space mission users may use command or notification pipelines to initiate the necessary activities to execute space data flows to their observatories, as described in Fig. 11 and Fig. 14 respectively. The queue arrival process for a transient science scenario can therefore be characterized by two factors: the mean occurrence rate of a transient scientific event and the number of associated space data flows it initiates. The number of associated space data flows (i.e. the batch size) initiated by occurrence of the transient scientific event depends on both the scenario design and the space network data flow implementation.

Transient scientific events commonly occur following a Poisson arrival process [38] [15]. The Poisson probability mass function and cumulative distribution function for a class of transient scientific events (or, more precisely, the functions describing the probabilities of their detection) can be used to estimate the frequency and number of expected occurrences that a transient science observation scenario will be triggered over a time interval.

The number of space data flows initiated by the transient event can be deterministic or a random variable depending on the scenario design. For example, three observatories could always need space data network services when a gravitational wave event is detected yet have differing thresholds for triggering space data flows in response to Gamma-ray burst detections. In this case, the probability functions of the access request batch size should be specified in the scenario design. For unspecified, emergent or ad hoc transient science scenarios, the statistical characteristics of the batch size may be empirically derived and used to inform operational network traffic load-balancing approaches.

A generalized framework for modeling a simple scenario involving a deterministic ground-to-space multipoint data flow may be expanded to investigate more complex scenarios involving

bi-directional sequences of multipoint data flows over time. In such a multi-tiered scenario, one or more of the first-tier data destination observatories may publish new information about the transient source, initiating a second-tier set of multipoint space data flows, and so on. The probability of initiating a conditional branch in a multi-tiered scenario should be specified by an additional random variable for each branch.

The probability functions for the occurrence process of a transient scientific event and its associated demand for space data flows in a scenario serve as inputs to the queueing model's servicing processes. The relationships between the queue arrival processes and servicing processes determine the system timeliness, capacity and throughput outcomes. Characterization of a scenario's queue arrival processes can be helpful to mission and network planners by providing a basis for first-order scenario feasibility screening using established servicing capabilities and constraints of candidate space data network solutions.

Queue Servicing Process Modeling Guidelines

Next, we consider the space data network access request servicing processes. The access request servicing processes are strongly influenced by user and network service provider design factors, design implementations and constraints. Accordingly, understanding the relationships and design trades that influence servicing processes within a transient science space system can be used to inform opportunities to conduct short-term transient science pathfinder scenarios within the constraints of existing user and space data network implementations as well as to identify longer-term mechanisms to enhance system capabilities.

User segment design factors that impact the servicing processes include ground and flight software component synchronization delays, the data volume of the data product to be

transported, whether data contents are unique to the observatory (i.e. from a command pipeline) or common with other space-based observatories in the scenario (i.e. from a notification pipeline) and the observatory orbit parameters, which describe its position and velocity state vectors as a function of time. If users employ manual procedures to perform the command pipeline activities, these delays should also be included into overall scenario latency budgets.

Network service provider segment design factors that impact servicing processes include the number and geographic coordinates or orbit parameters of the servicing nodes (i.e. direct-to-Earth ground stations or space relays), pre-service and inter-service configuration setup and teardown delays, discrete software, hardware or protocol-imposed data rate limits, and the number of simultaneous space link service instances available per servicing node.

The interfaces between user and network segments also impact servicing processes. Traditional space mission analysis and design practices focus much attention on the design of the service execution interface. This includes user and network design choices for mutual compatibility (e.g. spectrum band, frequencies, and space telecommunications standards), and satisfaction of tracking and link closure criteria (e.g. pointing, range, velocity, doppler offsets, signal transmit power and receiver sensitivities, data rates, link margin thresholds, acceptable bit error rates, etc.). The ephemerides of user observatories in relation to network service provider nodes can also influence delays due to light travel time and the traffic loading, as there may be periods when multiple observatories are in view of only one servicing node. Important service management interface design factors include the implementation and associated delays of the automated processing activities allocated to command and notification pipelines (described

respectively in Fig. 11 and Fig. 14), the space data network traffic loading and utilization patterns for other users, and the scheduling priority among users internal and external to the scenario.

There are additional factors that affect the servicing processes related to the operational assumptions and capabilities of the user and network service provider segments. For example, users may require synchronous bi-directional data flows to assure delivery of commands or notifications sent to the observatory or adopt a one-way commanding or notification approach. Alternative security and data transport protocols could also impact the servicing processes. For example, NASA's current space data networks implement circuit-switched data flows through bent-pipe space relays or directly with terrestrial ground stations. However, delay and disruption tolerant networking (DTN) protocols will enable more advanced types of network control functions and packet-switched data flows in space data networks, including multi-hop routing, in-transit data storage and queueing, data multicasting, and packet-level traffic prioritization capabilities. In addition, DTN protocols may allow for service management metadata to be included within the user data stream. This would allow space data network service providers to perform traffic management in a similar fashion to terrestrial IP networks and remove the need for the "out-of-band" service management activities depicted in the command and notification pipeline activity flows (Fig. 11 and Fig. 14 respectively). The impacts of alternative network security and DTN protocols on the system servicing processes can be explored using software-in-the-loop simulations within an orbit propagation and space communications link modeling environment.

System Suitability Metrics

System suitability metrics are quantifiable features of a system that can inform the likelihood and degree of satisfaction of the system objectives. Suitability metrics for evaluating space data network solutions for transient science space systems are provided in Table 11.

Table 11: Space data network suitability metrics for transient science space system design studies.

Suitability Metric Category	System Suitability Metric Definition	Notation
Timeliness Metrics	Service start wait time for the i^{th} observatory in the scenario	$SSWT_i$
	Time to complete scenario instance data flows	*TTC*
Capacity Metrics	Number of space data network service instances per scenario instance	*NSI*
	Long-term space data network utilization imposed by the scenario	*NU*
Throughput Metrics	Scenario instance data volume throughput	*DV*
	Effective scenario servicing data rate	*EDR*
	Long-term scenario data volume throughput	*LTDV*

A hypothetical design study illustrating a transient science space system queueing model implementation and use of the suitability metrics to evaluate alternative scenario and space data network solutions will be presented in Section 3.

Multipoint Space Data Flow Design & Evaluation Method Definition

A transient science observation scenario is instantiated by time-sensitive data flows that orchestrate space data network access and data transport activities involving one or more space-based observatories. The suitability of alternative implementation solutions for achieving the scenario objectives can be explored quantitatively through modeling and simulation of the transient science space system's queue arrival and servicing processes. System suitability metrics

provide indicators for mission planners and scientists as to the overall scenario feasibility and relative merits of alternative scenario data flow implementations. An iterative data flow design method for transient science space system design studies is illustrated in Fig. 15.

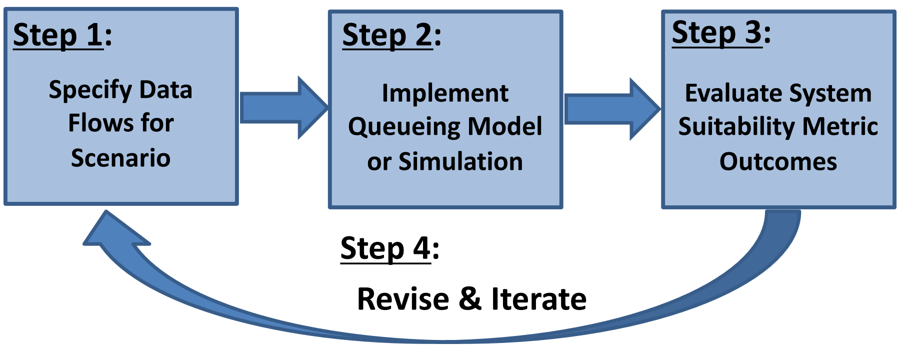


Fig. 15: Data flow design method for transient science space system design studies.

The iterative scenario data flow definition, analysis and evaluation method depicted in Fig. 15 can be further decomposed and implemented according to the steps below.

1. Specify the data flow characteristics of the scenario.
    a. Specify data types involved in the scenario.
        i. Multipoint Event Notification Data (e.g. common alerts, operating status or other situational awareness information sent to one or more destination)
        ii. Science Mission Data (e.g. instrument #1 data, instrument #2 data, etc.)
        iii. Engineering Data (e.g. spacecraft commands, telemetry, etc.)
    b. Specify source/destination data flow relationships among user elements, including any sequencing requirements or conditional branches in the scenario.
        i. Single source to single destination (One-to-One) e.g. a notification of a volcanic eruption event sent to a single specified Earth science space user MOC or

observatory [7]. The data contents for One-to-One flows may be non-standard since they are intended for a single destination.

ii. Single source to multiple destinations (One-to-Many) e.g. a notification of gravitational wave event multicast to subscribed astrophysics ground and space-based observatories or other applications where data products may be replicated and multicast or published for dissemination to a set of pre-specified or subscribed users. The data contents for One-to-Many flows provide opportunities for flexibility and efficiencies to network service providers if they are standardized among the user destinations for a given multi-cast data subject or topic area.

iii. Single source to all destinations (One-to-All) e.g. a network advisory message indicating a service outage, a solar energetic particle health and safety alert, or other applications where data products may be replicated and broadcast to all users. The data contents for One-to-All flows provide opportunities for flexibility and efficiencies to network service providers if they are standardized globally across all user destinations.

c. Specify the timeliness goals or threshold requirements for the data flows.

d. Specify the data volume of data products in the scenario data flows.

e. Specify the occurrence rate or probability distribution characteristics of the triggering scientific event, as well as the probabilities for any conditional data flow branches in the scenario.

2. Implement system queueing behavioral model or simulation.

a. Identify and characterize user-network service provider node topology factors (e.g., node coverage, loading, compatibility, range, and link closure satisfaction criteria)
b. Identify and characterize service management factors (e.g., service access request and provisioning processes)
c. Identify and characterize service execution factors (e.g., signal acquisition, synchronization latencies, data rates).
d. Formulate and execute a quantitative model or simulation that predicts outcomes for the system suitability metrics.

3. Evaluate outcomes for the system suitability metrics.
4. Revise and iterate scenario or data flow implementation alternatives.

Summary

Queueing theory provides a well-established mathematical framework for the design and evaluation of transient science space systems. Many user and network service provider design factors can impact the outcomes of the system suitability metrics. In some cases, simplifying assumptions can be made to adequately model the queueing processes and determine the feasibility of a scenario implementation by a candidate space data network solution. In other cases, the relative influence of the myriad design factors on the outcomes of the system suitability metrics can be more accurately and precisely determined using Design-of-Experiments and Monte Carlo techniques.

Transient Science Space System Design Study

A transient science space system is comprised of observatories whose operations are connected by purposefully designed time-sensitive data flows. The objective of Section 3 (this section) is to apply the command pipeline and notification pipeline functional reference architectures developed in Section 1 and the multipoint space data flow design and evaluation method developed in Section 2 to an example transient science space system design study. Four multipoint space data flow implementations using NASA's TDRS space relay data network are examined in the context of a transient science observation scenario.

Scenario Definition

Three notional space-based astrophysical observatories are in the planning phase of development and have engaged NASA's Near Space Network planners to design and evaluate space data flow implementations for their missions. Each space-based observatory is owned and operated by different entities (e.g. by NASA itself, by a NASA international partner and by a NASA academic partner) and will be controlled by separate operations centers. The observatories will perform independent pre-specified survey observations of differing astrophysical sources during nominal operations. Additionally, mission planners seek to understand the feasibility and design options for a transient science observation scenario involving the three space-based observatories. The data flows for nominal operations will not be analyzed in this example design study for brevity.

The transient science observation scenario is initiated by ground-based detection of gravitational-waves associated with the in-spiraling and merger of two neutron stars. The objective of the scenario is to perform follow-up observations of electromagnetic emissions using

the three space-based observatories as soon as possible following notification from the ground-based observatories. Information about the neutron star merger transient event is disseminated from the ground-based observatories to subscribed users in the science community via the Gamma-ray Burst Coordinates Network (GCN) transient science messaging application.

Due to launch manifest and logistical considerations, the three space-based observatories will be deployed from the International Space Station as a string-of-pearls constellation with the following characteristics: a circular orbital plane at 400 kilometers altitude, 52 degrees inclination with a 15 second separation time between the observatories (i.e., each observatory orbits with the same longitude of the ascending node and maintains an orbital phasing offset in true anomaly of approximately 2 degrees between each observatory).

An operational view of the end-to-end transient astrophysical processes, observables and transient science observation scenario is illustrated below in Fig. 16.

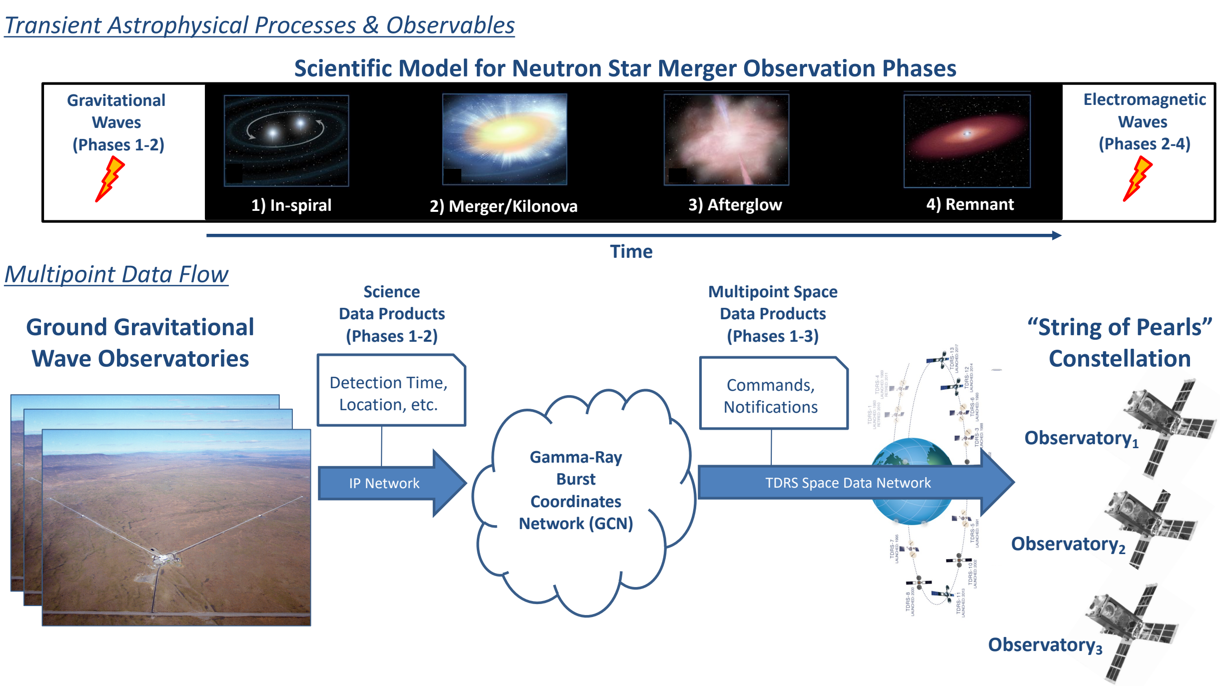


Fig. 16: Operational view of astrophysical processes and a transient science observation scenario for neutron star merger events.

Design Study Scope

The Near Space Network oversees NASA's network planning and operations for missions venturing out to two million kilometers from Earth. The Near Space Network is comprised of commercial, international and NASA owned and operated direct-to-Earth ground stations as well as the NASA owned TDRS space relay data network. User command and science data associated with pre-planned nominal mission operations are well-suited for servicing by direct-to-Earth ground stations and can be designed using traditional methods. However, transient science observation scenarios pose several challenges for mission and network planners due to their random occurrence, their need for timely follow-up observations and the resulting amplification of demand and traffic for space data network services. The command and notification pipeline reference architectures and the multipoint space data flow design and evaluation method provide new tools for mission and network planners to address these challenges. The TDRS space relay data network is amenable to the needs of some simple transient science observation scenarios for observatories in low-Earth orbit because of its coverage, service management and service execution capabilities. As a result, the scope of network solutions considered by the design study will be limited to multipoint data flow implementations using the TDRS space data network.

TDRS Coverage

TDRS provides global communications coverage for observatories in low-Earth orbit. Global coverage is achieved by stationing at least three geosynchronous TDRS satellites (i.e. TDRS nodes) in overlapping service coverage regions, known as the Atlantic Ocean Region, Pacific Ocean Region and Indian Ocean Region. Geosynchronous satellites have circular orbits with

altitudes of approximately 36,000 kilometers above the Earth's equator and complete one orbit approximately every 24 hours. Each TDRS node could communicate with a low-Earth orbit satellite for approximately 60% of its orbital period with an irreducible light travel time delay of approximately one quarter of one second.

NASA's low-Earth orbiting space-based observatories commonly operate in nearly circular orbits at altitudes ranging from 400 to 850 kilometers, and at inclinations from equatorial to polar sun-synchronous, roughly 0 to 100 degrees. The orbital period of a low-Earth orbit satellite is approximately one orbit every 90 minutes. As a result, low-Earth orbiting satellites will transit one of the three TDRS service coverage regions in approximately 60 minutes. However, the relative positions of low-Earth orbiting satellites will generally vary due to differences in orbital altitude, inclination, orbit maintenance needs, maneuver timing, and perturbations due to the oblateness of the Earth, atmospheric drag and solar radiation pressure among other factors. As a result, the number of low-Earth orbit satellites within a TDRS service coverage region (or in view of a TDRS node) is a random variable whose characteristics can be empirically estimated for a defined time period using powerful orbit propagation and communications link simulation tools such as Systems Tool Kit.

TDRS Service Management & Service Execution Capabilities

The TDRS Demand Access System provides transient science observatories, such as Swift, with timely access to space-to-ground (i.e., return) data transport services. Timely space-to-ground data transport services for multiple simultaneous users are achieved using multiple-access beamformed links and pre-allocated network resources. Specifically, TDRS nodes and associated strings of ground signal processing and data handling equipment are allocated to

continuously track and handover service responsibilities for the observatory as it transits the coverage region of each TDRS node [32].[8] Timely space-to-ground data transport services allow space-based observatories to initiate multi-observatory transient science observation scenarios through the GCN science messaging application.

The TDRS space data network does not presently provide users with continuous access to ground-to-space (i.e., forward) data transport services. However, an experimental TDRS broadcast link was demonstrated in 2016 and could provide multiple simultaneous users with continuous ground-to-space data transport services [36]. The broadcast link could be implemented using the forward link beamformers and antenna array on TDRS nodes.[9] Alternatively, the as-built TDRS service management interface allows users to invoke event-driven ground-to-space data transport services using traditional beamformed forward links.[10] The latencies associated with event-driven ground-to-space data transport services have been well characterized in Chapter 3 and in Roberts et. al. [46].

Design Study Point of Departure

The ultimate objective of the multipoint data flow design and evaluation method is to facilitate stakeholder consensus on a network solution for the transient science observation

---

[8] The TDRS constellation is currently comprised of three third generation nodes, three second generation nodes and four first generation nodes. Demand Access System services are available from first and third generation nodes only due to implementation constraints in the second-generation nodes.

[9] Presently all 10 operational TDRS nodes can support one forward service instance using the beamformed antenna array. Each second and third generation TDRS node could support up to two simultaneous forward service instances with ground segment upgrades. As a result, Near Space Network planners have operational and engineering development options available to implement a TDRS broadcast service without adverse impacts to pre-scheduled users.

[10] Beamformers direct the link to a space-based observatory and track it during a service execution period. TDRS link design characteristics and service configuration parameters typically vary for different users. A TDRS broadcast link has standardized signal characteristics and is directed to encompass the node's coverage region at the expense of signal power density and hence achievable data rate.

scenario. As a point of departure, the design study will develop executable models to evaluate the relative outcomes on the system suitability metrics for two implementations of the multipoint data flows in the scenario: 1) a command pipeline and TDRS node servicing each observatory sequentially using beamformed links and 2) a notification pipeline and TDRS node servicing each observatory in parallel using a broadcast link. Evaluation of the point-of-departure results may lead to revision and iteration of the scenario and data flow design factors until a consensus network solution is achieved or the scenario is determined infeasible with the considered network implementations.

Physical views of the command and notification pipeline data flow implementations for the point-of-departure network implementations are provided by Fig. 17 and developed in the following section.

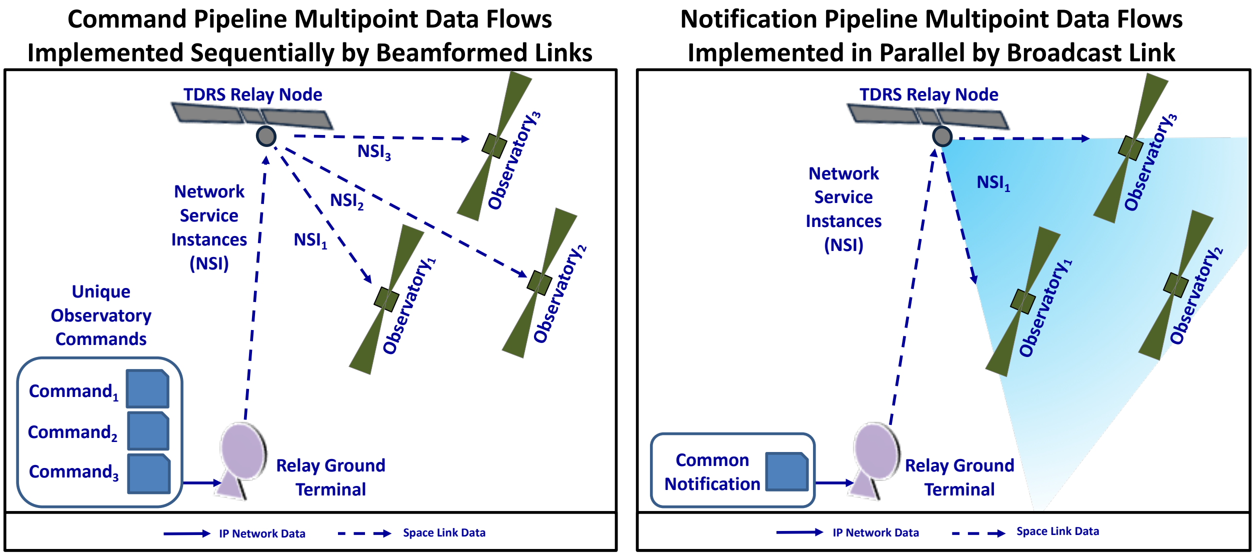


Fig. 17: Physical view of command pipeline (left) and notification pipeline (right) multipoint data flows implemented by TDRS.

Multipoint Space Data Flow Design and Evaluation Method Application

The multipoint space data flow design and evaluation method defined in Section 2 is applied to the defined transient science observation scenario within the scope of the design study.

Step 1: Specify the data flow characteristics of the scenario.

The data flow characteristics of a scenario are specified in a five-part sequence, described below.

Step 1a: Specify data types of the scenario.

This scenario may involve commands (engineering data types) or common messages describing attributes of the transient source such as time of detection and celestial location (notification data types). For simplicity, engineering telemetry from the space-based observatory to the ground indicating receipt of the command or notification data onboard the space-based observatory is not required.

Step 1b: Specify source/destination data flow relationships among user mission elements, including any sequencing requirements or conditional branches in the scenario.

Transient science event information received from the science messaging application initiates an activity sequence to implement multipoint ground-to-space data flows. There are three space-based observatories in this scenario. There are no data delivery order sequencing requirements among the observatories nor are there conditional data flow branches in this scenario. Command data originates from separate user command pipelines instantiated at user mission control elements and ends at the respective user space-based observatories. Notification data originates at the network notification pipeline and ends at the user space-based observatories.

Step 1c: Specify the timeliness goals or threshold requirements for the data flows.

The command or notification data flows shall be delivered as soon as possible to the space-based observatories and must be delivered no later than 10 minutes (600 seconds) following initiation of service access processes for space data network services.

Step 1d: Specify the data product volumes in the scenario data flows.

Command data is unique to each space-based observatory and for this scenario has a data volume of 800 bits per observatory command. Notification data is common to all space-based observatories involved in the scenario and has a data volume of 800 bits per notification message in this scenario. Information security, error-correction encoding, and other data transport protocol overheads are assumed equivalent for command and notification data and are ignored.

Step 1e: Specify the occurrence rate or probability distribution characteristics of the triggering scientific event as well as the probabilities for any conditional data flow branches in the scenario.

It is estimated that a detectible neutron star merger event occurs following a memoryless random occurrence process with a mean occurrence rate of 5 events per day. There are no conditional data flow branches in this scenario.

Step 2: Implement system queueing behavioral model or simulation.

The queueing behavioral model and simulation is implemented by identifying and mathematically relating the arrival and servicing factors described below. Communications coverage simulations and empirical experiments with the as-built TDRS systems provide the basis for estimating the queue arrival and servicing factors associated with the command and notification pipeline implementations.

Step 2a: Identify and characterize user-network service provider node topology.

The user-network service provider node topology characteristics can contribute latency in the data flow due to light travel time, gaps in coverage and gaps in satisfaction of the link closure criteria between the observatories and TDRS nodes. In addition, node topology characteristics support quantification of the impacts of the scenario's correlated user service demand. Specifically, scenario node topology simulations can provide the set of possible TDRS node servicing opportunities for the observatory. When combined with other design factors that affect data flow implementations, this information leads to the scenario's queue arrival distribution per TDRS node (or, stated alternatively, each TDRS node's expected loading). This distribution is characterized by the number of space-based observatories in the scenario to be serviced by a TDRS node within a defined time period.

The orbital elements of the user and network service provider nodes allow the relative positions of the nodes to be determined over time. The space-based observatories are to be launched from the International Space Station into a constellation consisting of a single circular orbital plane at 400 kilometers altitude, 52 degrees inclination and are phase separated by 15 seconds. The TDRS nodes are in a circular geosynchronous orbit at roughly 36,000 kilometers altitude, with small but varied inclinations, and are phase separated into three service coverage regions around the Earth. As described previously, all satellites vary in their orbital altitude, inclination and phasing parameters due to mission needs, natural perturbations and operational decisions. The North American Defense Command publishes a public catalog of precise orbital elements for satellites and other objects in space that are derived from regular observations. A commercial aerospace modeling and simulation platform known as Systems Tool Kit (STK) was used to ingest publicly available orbital elements for the TDRS constellation, propagate the orbits of all satellites involved in the scenario over a defined time period, and compute the overall coverage and individual TDRS node loading results. A single TDRS node from each service coverage region was simulated. The orbits were propagated using Satellite Took Kit's J4 propagator algorithm which includes perturbations on the orbits over time because of the Earth's oblateness. A visualization of the ground track for TDRS nodes and observatories from the simulation is provided in Fig. 18.

Fig. 18: Visualization of the ground track for TDRS nodes and observatories from the STK simulation.

Verification procedures were performed to investigate the sensitivity of results to the simulation timestep and analysis period (i.e. the simulated duration of time over which coverage and TDRS node loading statistics are calculated). After several iterations, a time step of 60 seconds over a simulated period of 35 days was found to give replicable results without losing much fidelity as compared to smaller timesteps and longer simulated periods. For example, as compared to a run with a timestep of 1 second and a simulation period of 180 days, the maximum, minimum and median coverage statistics for a single TDRS node and user satellite did not change, while the mean and standard deviation varied only slightly at 0.50% and 1.70% respectively.

The simulation results verify global coverage of the user constellation with the specified TDRS nodes. There is substantial overlap in coverage among the three TDRS nodes, with each TDRS node covering approximately 60% of the observatories' orbits. It is assumed that the TDRS nodes operate at 100% availability, with no planned or unplanned outages due to maintenance activities or reliability issues.[11] Therefore, there is no latency in the scenario due to gaps in

[11] In practice, availability approaching 100% is achieved through constellation-level load balancing and redundancies.

coverage. The simulation results also verify the maximum total latency due to light travel time is approximately one quarter of one second.[12]

A suitable user terminal, such as the Vulcan Wireless NSR-SDR-S/S, was selected for each observatory to satisfy compatibility and link closure criteria for use of the services of the TDRS constellation [52]. Adequate link margins ensure there is no latency induced by signal dropouts during service execution.

The simulation results indicate that despite substantial coverage overlap by the three TDRS nodes, there are substantial periods when the observatories are concentrated within the coverage area of a single TDRS node due to their small orbital separation distance. As a result, the queueing model can be simplified by analyzing the outcomes for when a single TDRS node must service all three observatories in its coverage region without relying on the coverage overlap from TDRS nodes in other regions. This simplification captures the correlated demand and traffic amplification aspects of the scenario that are not presently analyzed by current space data network planning methods. However, it will produce results that predict somewhat worse outcomes than the more complicated case where coverage overlaps from more than one TDRS node are considered.

Step 2b: Identify and characterize service management factors.

The network service provider service management activities differ in the command and notification pipeline reference architectures. For command pipelines, notification from the science messaging application about a transient scientific event initiates the creation of an access

[12] Total light travel time is computed based on the range from the TDRS ground terminal to the TDRS servicing node in geosynchronous orbit and from the TDRS servicing node to the space-based observatory.

request within the operations center of each space-based observatory. This process results in a batch of service access requests arriving to the space data network's service management interface. Notification pipelines provide a subscription-based transient science data content delivery service to space-based observatories. They are implemented by the network service provider as an extension to traditional space data network service management functions. Notification pipelines interface directly with terrestrial science messaging applications. User mission controllers pre-specify their space data network access preferences based on the attributes of the transient event message received directly by the notification pipeline from the science messaging application. The notification pipeline orchestrates space data network access activities to deliver standardized space notification messages to observatories in accordance with users' preferences and scenario objectives.

Command Pipeline Implementation

A command pipeline automates the follow-up decision logic, observatory command-build and TDRS service access request activities. It is assumed that the command pipeline instantiated at each of the mission operations centers would have similar decision processing and command build latencies, and that variations in the data transport time over terrestrial Internet Protocol networks from the different control centers is negligible. A command pipeline automates the build of service access requests to the TDRS service management and network control subsystems using an External Processing System. The service access request build latency using an External Processing System has been estimated as less than 5 seconds.

The three service access requests arrive simultaneously at the TDRS service management interface and are queued for processing. To estimate the request processing time, a total of 30

service access requests were submitted to the engineering test instance of the TDRS service management and network control subsystems. The median processing time per request was 2 seconds with a standard deviation of 7 seconds.

The TDRS service management and network control subsystems implement a minimum service start lead time following disposition of a service access request. Experimental results demonstrated that services could be successfully scheduled with as little as 6 minutes of lead time. Experimental results verified the minimum inter-user service lead time of 30 seconds. The minimum service start lead time and minimum inter-user service lead time were also experimentally verified to be fixed values imposed by TDRS network subsystems.

Event-driven access requests from command pipelines are fulfilled using unreserved periods on TDRS nodes and allocated on a first-come-first-served basis. As a point of departure for the design study, the probability of blocking due to pre-scheduled user service periods and the associated latencies are ignored for simplicity. Data flow implementations that satisfy the scenario timeliness requirements using simplified models should be re-analyzed with higher-fidelity models before stakeholders select this implementation as the preferred network solution. Scenarios that cannot be satisfied despite ignoring the impacts of blocking are likely infeasible with this command pipeline implementation.

Notification Pipeline Implementation

A notification pipeline implemented through a continuous TDRS broadcast service does not yet exist. However, an experimental TDRS broadcast link has been demonstrated and provides a basis for estimating the performance attributes of the notification pipeline implementation for this scenario [36].

A broadcast service has a one source-to-all destinations data flow relationship. Because the notification pipeline disseminates a common standard message to all space-based observatories via a continuous transmission link, the service management and associated network control functions are greatly simplified from the general notification pipeline reference architecture activity flow shown in Fig. 14. The general notification pipeline reference architecture could be applied in federated multi-network service provider topologies (i.e., where the notification pipeline interfaces with more than one service management element) or to disseminate notification messages to multiple space-based observatories within the same beamwidth of a beamformed link if a continuous TDRS broadcast service is not implemented.

For the TDRS broadcast service implementation, the notification pipeline receives messages from the terrestrial science messaging application, converts them to a standard space data format, such as the Asynchronous Message Service (CCSDS 735.1-B-1), and disseminates them to the TDRS ground terminal for broadcast to space-based observatories in accordance with users' pre-defined content delivery preferences. This process is fully automated and is estimated to take less than one second.

Step 2c: Identify and characterize service execution factors.

The network service provider's service execution activities are the same for the command pipeline and notification pipeline reference architectures. However, differences that affect the system suitability metric outcomes arise in the implementation of the scenario's service execution activities using beamformed links as compared to broadcast links.

Command Pipeline Implementation

Command pipelines generate unique observatory commands for each user in a transient science observation scenario. Space-based observatories typically have differing service execution link parameters and configuration settings. As a result, three unique beamformed links, each a TDRS network service instance, are required to implement the multipoint data flow for this scenario. At the service execution start time, a signal is directed from the TDRS node to the user observatory, implementing the final link in the data flow.[13] At the maximum TDRS forward data rate of 300,000 bits per second, the time required to transport the specified command data volume to an observatory in this scenario is less than 0.003 seconds. However, there are implementation specific latencies associated with the observatory's acquisition of the TDRS signal and synchronization of the user flight and ground systems. The latency to transport the command data volume is much less than the combined acquisition and synchronization latencies experienced during service execution. For example, a typical event-driven service execution period for the Swift mission is two minutes to allow adequate margin for the acquisition and synchronization processes, yet once the link has been established data transport from the Swift control center to the observatory takes less than 1 second. The minimum allowable TDRS service execution period was experimentally verified to be 60 seconds. For the hypothetical design study, it is assumed that the user signal acquisition, software synchronization and data transport processes can be completed within the fixed 60 second TDRS minimum

[13] A command pipeline data flow originates at the user's mission operations centers. The data flow passes over a terrestrial Internet Protocol network to the TDRS terrestrial network gateway. From the TDRS gateway, the data passes to a TDRS ground terminal and into space to the TDRS node where it is then forwarded from the TDRS node to the observatory by the beamformed link.

service duration. As a result, beamformed link service instances have a deterministic execution duration of 60 seconds.

Execution of the first service instance occurs following the minimum service lead time interval. Following completion of the first service instance, an inter-user lead time of 30 seconds occurs prior to initiation of the second service instance. Similarly, following completion of the second service instance, an inter-user lead time of 30 seconds occurs prior to initiation of the third service instance. The scenario data flows are complete at the end of the third service execution period.

Notification Pipeline Implementation

The notification pipeline in this study is implemented by a TDRS broadcast service that disseminates common standard space messaging data to all observatories using a single continuous service instance per TDRS node. The TDRS broadcast service is implemented by directing the antenna array to radiate with a wide beamwidth toward the Earth. Use of the broadcast service requires all observatories to implement subsystems that receive, identify, process and respond to common notifications encoded in the broadcast signal. This eliminates the latencies associated with the minimum service start lead time, signal acquisition and synchronization time and the inter-user lead time associated with discrete beamformed link implementations.

In July 2016, Heckler et al. demonstrated the feasibility of implementing a TDRS broadcast service [36]. Using the TDRS-12 node, four elements of the TDRS forward antenna array were placed in broadcast configuration. Signal power measurements were made at the ground and onboard the International Space Station. These measurements validated the modeling and

simulation performance predictions. A TDRS broadcast service user terminal design suitable for the resource limitations of a small satellite platform was also identified. The experimental results showed that a TDRS broadcast link can support data rates of 1024 bits per second with robust margins.

As a point of departure for this study, the full bandwidth of the broadcast link is assumed to be available for allocation to servicing the scenario's data flow needs. Scenarios that cannot be satisfied despite full allocation of the TDRS broadcast link bandwidth are likely infeasible with this notification pipeline implementation and may require a direct user cross-link topology.

Step 2d: Formulate and execute a quantitative model or simulation that predicts outcomes for the system suitability metrics.

The degree of fidelity and complexity of the quantitative model or simulation needed to make valid and meaningful distinctions among implementation options in a design study depends on many factors. Models or simulations for screening the feasibility of candidate solutions in the conceptual development phase may simplify some details that would be important to include in models used for making network service commitments. Careful application of quantitative reasoning and engineering judgement are necessary to substantiate and justify modeling and simulation implementation choices.

Step 1 and Step 2 of the multipoint data flow design and evaluation method identified several factors that can impact the queue arrival and servicing processes for the scenario. These factors are formulated into quantitative models for the point of departure data flow implementations to predictively quantify the outcomes of the system suitability metrics introduced in Table 11. The models developed subsequently are intended to illustrate a suitable

level of fidelity for initial feasibility screening and relative comparison of the implementation options for the transient science observation scenario. As a result, some factors judged to have small impacts on the suitability metric outcomes or common to all implementations (e.g., light travel time) are not modeled.

Command Pipeline Quantitative Model Implementation

A summary of the queue arrival and servicing design factors and their characteristics for the command pipeline implementation identified in Steps 1 and 2 of the data flow design and evaluation method are provided in Table 12 below.

Table 12: Summary of data flow design factors for the command pipeline implementation.

Factor	**Arrival/Service**	**Step**	**Fixed or Random**	**Value**	**Basis**
Number of Space-Based Observatories	Arrival	1b	Fixed	3	Specified by Scenario Design
Occurrence Rate of Transient Science Event	Arrival	1e	Random	5 per day (mean)	Specified by Scenario Design
Data Volume	Service	1d	Fixed	800 bits	Specified by Scenario Design
Light Travel Time	Service	2a	Fixed	< 1 second	Simulation
Coverage Gaps	Service	2a	Fixed	0 seconds	Simulation
Link Dropouts	Service	2a	Fixed	0 seconds	Documentation
Service Access Request Build	Service	2b	Fixed	5 seconds	Estimate
Service Access Request Processing	Service	2b	Random	2 seconds (median)	Experimental

Minimum Service Start Lead Time	Service	2b	Fixed	360 seconds	Experimental
Inter-User Lead Time	Service	2b	Fixed	30 seconds	Experimental
Blocking	Service	2c	Fixed	0 seconds	Specified in Implementation
Data Rate	Service	2c	Fixed	300,000 bits per second	Documentation
Minimum Service Execution Period	Service	2c	Fixed	60 seconds	Experimental
Number of Service Instances Required to Execute Scenario	Service	2c	Fixed	3	Specified by Implementation

Quantitative models to predict outcomes for each of the system suitability metrics for the scenario's command pipeline data flow implementation are developed below. These models are generally applicable when all observatories in a scenario using command pipelines must be serviced by a single TDRS node with beamformed forward links.

Timeliness Modeling

There are two system suitability metrics for timeliness. The first, service start wait time for the i^{th} observatory, $SSWT_i$, measures the queue wait time for each user to begin its servicing by the TDRS space data network. It is computed as the expected time interval from initiation of the service access process until the service start time for the i^{th} observatory in a scenario with N total observatories.

Three TDRS queue servicing factors contribute to the service start wait time: the minimum service start lead time, *Min_Lead*, the minimum service execution time per user, *Min_SE*, and the inter-user lead time, $Inter\_User_{(i-1,i)}$, which is a transitional configuration teardown and setup period that occurs between users serviced sequentially by a TDRS node. The service start time for each observatory in the scenario can be calculated by Eq. 9.

$$SSWT_i = \text{Min\_Lead} + \sum_{i-1}^{N} \left(Min\_SE_{(i-1)} + Inter\_User_{(i-1,i)} \right) \tag{9}$$

The second timeliness metric is the time to complete the scenario service instance, *TTC*. *TTC* measures the expected time interval following initiation of the service access process until the final service execution period ends, completing the data flows in the scenario instance. The time to complete the scenario service instance can be calculated by Eq. 10.

$$TTC = SSWT_N + Min\_SE_N \tag{10}$$

Capacity Modeling

There are two system suitability metrics for capacity. First is the number of space data network service instances required per scenario instance, *NSI*. For the simple deterministic data flows of this scenario, *NSI*, is specified in the implementation. It is equal to the total number of space-based observatories in the scenario, *N*, provided in Eq. 11.

$$NSI = N \tag{11}$$

Second, the long-term space data network utilization of a scenario, *NU*, measures the expected time the network is busy servicing the scenario divided by the total network capacity available for all users. The network utilization metric provides a basis for evaluating the impact

of the long-term space data network loading that will be imposed on the network service provider for a scenario implementation.

The expected time the space data network is busy servicing the scenario is the product of the number of times the scenario is expected to occur multiplied by the expected servicing time required to complete scenario instance. The number of scenario instances that are expected to occur is based on the expected occurrence rate of the transient scientific event, *Event_Rate*. The time to complete a scenario instance, *TTC*, is provided in Eq. 10. The *TTC* metric is inclusive of the minimum service start lead time, *Min_Lead*. Since TDRS nodes are not busy servicing the scenario during the minimum service lead time, the *Min_Lead* factor is subtracted from *TTC* to obtain the network servicing time for the scenario instance.[14]

The total network capacity is computed as the product of the number of individual TDRS nodes providing services for the scenario, *Num_Servicing_Nodes* and their availability, *Node_Availability*.[15] The evaluation period is the duration of time over which outcomes for the system suitability metrics are evaluated and is carefully chosen to adequately characterize the long-term effects of randomness in the system behavior. The 35-day simulation period, *Simulation_Period*, was found to adequately characterize TDRS node coverage and loading variations using model verification procedures in Step 2a of the data flow design method. The

---

[14] The TDRS service management interface imposes the minimum service start lead time, *Min_Lead*, constraint on all new service access requests whether they are submitted weeks in advance through the nominal scheduling process or generated from command pipelines in response to a transient scientific event. Users who have a pre-reserved service access period on a TDRS node that overlaps with the minimum lead time of an event-driven access request from a command pipeline are not in conflict, provided the pre-scheduled service access period ends prior to the end of the minimum lead period of the event-driven access request. More precisely, to avoid a scheduling conflict, the pre-scheduled service period must end, and the inter-user lead time must occur prior to the service start time of the event-driven service period. Wait times due to blocking are not considered in this design study for simplicity.

[15] More precisely, the total network capacity is calculated on a servicing resource basis. Since presently each TDRS node can provide only one beamformed forward service instance at a time, the number of servicing nodes is used.

long-term mean occurrence rate of the transient scientific event, *Event_Rate*, is specified in the scenario. For more sophisticated analyses, the distribution characteristics of the scientific event could be specified based on empirical scientific datasets or theoretical models and the effects of the occurrence rate uncertainties could be analyzed using simulation and statistical methods. Such methods would allow the estimation of confidence intervals for values of the affected system suitability metrics.

The long-term network utilization of a scenario is provided by Eq. 12.

$$U_{Scenario} = \frac{Event\_Rate * (TTC - Min\_Lead)}{Num\_Servicing\_Nodes \ * \ Availability * Simulation\_Period} \quad (12)$$

Throughput Modeling

There are three system suitability metrics for throughput. First, the total data volume transported by the space data network for a scenario instance, *DV*, is the sum of the command data volume sent by each of the command pipelines to their respective observatories, DV_i. The command data volume and number of user command pipelines are deterministic as specified in the scenario implementation and given by Eq. 13.

$$DV = \sum_{i=1}^{N} DV_i \quad (13)$$

Next, the effective scenario servicing data rate, *EDR*, provides a basis for evaluating the time efficiency of a scenario implementation. *EDR* is calculated as the ratio between the total data volume delivered for a scenario instance, *DV* (Eq. 13), and the scenario time to complete, *TTC* (Eq. 10). *EDR* is inclusive of queue servicing wait times in a scenario implementation that are not typically captured in traditional network feasibility analyses. *EDR* is calculated by Eq. 14.

$$EDR = \frac{DV}{TTC} \tag{14}$$

Finally, the long-term scenario data volume throughput, *LTDV*, quantifies the total data volume the space data network is expected to transport while a scenario is in operations. It is computed as the total data volume delivered for a scenario instance, *DV* (Eq. 13), and the occurrence rate of the transient scientific event, *Event_Rate*, provided in Eq. 15.

$$LTDV = \mathrm{DV} * \mathrm{Event\_Rate} \tag{15}$$

Notification Pipeline Quantitative Model Implementation

A summary of the queue arrival and servicing design factors and their characteristics for the notification pipeline implementation identified in Steps 1 and 2 of the data flow design and evaluation method are provided in Table 13 below.

Table 13: Summary of data flow design factors for the notification pipeline implementation.

Factor	Arrival/Service	Step	Fixed or Random	Value	Basis
Number of Space-Based Observatories	Arrival	1b	Fixed	3	Specified by Scenario Design
Occurrence Rate of Transient Science Event	Arrival	1e	Random	5 per day (mean)	Specified by Scenario Design (Observable Target)
Data Volume	Service	1d	Fixed	800 bits	Specified by Scenario Design
Light Travel Time	Service	2a	Fixed	< 1 second	Simulation
Coverage Gaps	Service	2a	Fixed	0 seconds	Simulation

Link Dropouts	Service	2a	Fixed	0 seconds	Experimental
Space Notification Format Processing	Service	2b	Fixed	< 1 second	Estimate
Data Rate	Service	2c	Fixed	1024 bits per second	Analytical
Number of Service Instances Required to Execute Scenario	Service	2c	Fixed	1	Specified by Implementation

Quantitative models to predict outcomes for each of the system suitability metrics for the scenario's notification pipeline data flow implementation are developed below. These models are generally applicable when all observatories in a scenario using notification pipelines must be serviced by a single TDRS node with a broadcast forward link and the full bandwidth of the broadcast link is allocated to servicing the scenario data flows.[16]

Timeliness Modeling

Several simplifying assumptions about the notification pipeline implementation are made for the service start wait time model. First, the notification message format and user preference processing latencies are assumed to be small and are ignored. All observatories are assumed to be connected to the TDRS broadcast link continuously. Small connectivity outages due to TDRS

[16] Although the design study scenario's node topology allows for a single broadcast link to service is data flows, three concurrent broadcast links must be provided by the TDRS constellation to perform handovers between TDRS service coverage regions and to ensure global coverage for scenarios involving observatories in more diverse orbits. Further investigation and model development work is necessary to study broadcast link bandwidth allocation trades for servicing the multipoint data flows of multiple concurrent scenarios involving general user orbits at the TDRS constellation level of analysis.

node handovers, channel perturbations or other sources are ignored. Finally, it is assumed that all observatories receive the standard notification message from the broadcast link simultaneously. As a result, the service start wait time for each observatory, $SSWT_i$, equals zero.

The scenario time to complete, *TTC*, occurs when the common standardized notification message is received by all observatories. Since all observatories are serviced by a single broadcast link and receive the same message in parallel, *TTC* is calculated as the notification data volume, *DV*, divided by the TDRS broadcast link data rate, *Broadcast_Data_Rate*, provided by Eq. 16.

$$TTC = \frac{DV}{Broadcast\_Data\_Rate} \tag{16}$$

Capacity Modeling

A single broadcast link network service instance, *NSI*, is required to transport the scenario data.

The long-term network utilization of a scenario, *NU*, is provided by Eq. 12 using the notification pipeline model for *TTC* (Eq.16) and a value of zero for the minimum service start lead time, *Min_Lead*, due to continuous connectivity of the broadcast implementation.

Throughput Modeling

The total data volume transported by the space data network for a scenario instance, *DV*, is specified by the scenario and equal to 800 bits.

The effective scenario servicing data rate, *EDR*, is calculated by Eq. 14 using the specified data volume, *DV*, and the scenario time to complete, *TTC (Eq. 16)*.

The long-term scenario data volume throughput, *LTDV*, is calculated by Eq. 15 using the specified notification pipeline value for total data volume delivered for a scenario instance, *DV*, and the and the occurrence rate of the transient scientific event, *Event_Rate*.

Step 3: Evaluate system suitability metrics.

The quantitative models developed in Step 2d of the multipoint data flow design and evaluation method were executed for the design study's point of departure scenario implementations to compute predictive outcomes for the system suitability metrics. The results are provided in Table 14: below.

Table 14: System suitability metric outcomes for command and notification pipeline scenario implementation options.

Suitability Metric Category	Metric	Command Pipeline Multipoint Data Flows Implemented Sequentially by Beamformed Links	Notification Pipeline Multipoint Data Flows Implemented in Parallel by Broadcast Link[17]	Units
Timeliness Metrics	$SSWT_i$	$SSWT_1$ = 360 $SSWT_2$ = 450 $SSWT_3$ = 540	0	Seconds
	TTC	600	0.78	Seconds
Capacity Metrics	*NSI*	3	1	Number
	NU	0.46 %	0.002 %	Percent
Throughput Metrics	*DV*	2,400	800	Bits
	EDR	3.2	1,024	Bits per Second
	LTDV	12,000	4,000	Bits per Day

[17] Several design factors were ignored in the broadcast link model implementation that would prohibit realization of such low absolute *TTC* and Network Utilization outcomes in practice. Nevertheless, accounting for queue servicing latencies identified in Table 13 but ignored in the quantitative models (e.g., light travel time and ground processing latencies to prepare and send space data), it is reasonable to expect at least one order of magnitude improvement in timeliness and utilization outcomes compared to the beamformed link command pipeline implementation due to the large fixed magnitude of the beamformed link start and inter-user lead times (600 and 30 seconds, respectively).

The timeliness goals and threshold requirements for the scenario data flows were specified in Step 1c of the data flow design method: "The command or notification data shall be delivered as soon as possible to the space-based observatories and must be delivered no later than 10 minutes (600 seconds) following initiation of service access processes for space data network services."

The *TTC* values for both command and notification pipelines satisfy the threshold requirement for data delivery under the stated assumptions of the models. However, the notification pipeline implementation achieves two orders of magnitude better delivery times for all observatories than the command pipeline implementation.[12] The network utilization for the scenario is more than 200 times less with the notification pipeline than for the command pipeline implementation.[12] The scenario instance data volume, *DV*, and long-term scenario data volume, *LTDV*, transported by the space data network is 3 times less with the notification pipeline. Finally, the effective data rate of the notification pipeline is over 300 times greater than that of the command pipeline implementation.

Although the models developed in Step 2d are relatively simple, the modeling results provide important information for system planners. First, the advantages of the notification pipeline implementation in the system suitability metrics are clear, whereas the command pipeline model results indicate little margin exists for meeting the 600 second *TTC* performance requirement. Careful attention to the command pipeline scenario and modeling assumptions are needed to improve confidence that the command pipeline implementation is a viable solution. For example, the TDRS constellation is currently comprised of 10 TDRS nodes. For service coverage regions with more than one TDRS node, the *TTC* for the command pipeline

implementation could be improved by parallelizing the observatory servicing tasks among multiple TDRS nodes. However, such an approach would require a mechanism for coordinating service access requests among the independent user command pipelines. The flexible scheduling parameters in a TDRS service access request message could provide such a mechanism, but it would need to be universally adopted by the users in the scenario. Alternatively, the command pipeline *TTC* could be improved by increasing the relative spacing among the observatories such that only one node is present in each TDRS service coverage region, enabling parallel TDRS servicing by beamformed links from three different TDRS nodes. However, this approach would not be viable if the increased spacing among observatories has unacceptable impacts to the science objectives.

The modeling results indicate that the command pipeline implementation bears a technical risk due to the predicted *TTC* performance and that the notification pipeline implementation is a superior solution on all system suitability metrics. However, factors such as achieving agreement on a common notification data standard and information security architecture among the scenario's independent mission users, implementing sufficient onboard autonomy for each of the observatories and establishing an engineering development project to implement the notification pipeline for the TDRS space data network may pose other kinds of risks to the viability of the notification pipeline implementation. In addition, both implementation options also carry unquantified uncertainties associated with resource contention from other transient science and traditional users. As a result of these risks and uncertainties, mission and network planners may need to revise and iterate the scenario design and implementation options.

Step 4: Revise and iterate scenario design or data flow implementation alternatives.

Based on the insights discussed in Step 3, two revisions to the implementation options are examined. The first revision pertains to the S*SWT* and *TTC* timeliness metrics for the command pipeline implementation. The second revision examines the system suitability metrics for a command and notification pipeline implementation using a shared broadcast service.

Revision 1: Command Pipeline Implemented with Parallel Servicing by Beamformed Links from Three TDRS nodes

The orbital parameters of the observatories were defined in Step 2a of the data flow design method. The first revision increases the spacing (i.e., difference in true anomaly) between the observatories from approximately two degrees separation to 120 degrees separation, allowing parallel servicing of each observatory by a beamformed link from the TDRS node in each of the three service coverage regions. A physical view of the revised scenario is provided in Fig. 19.

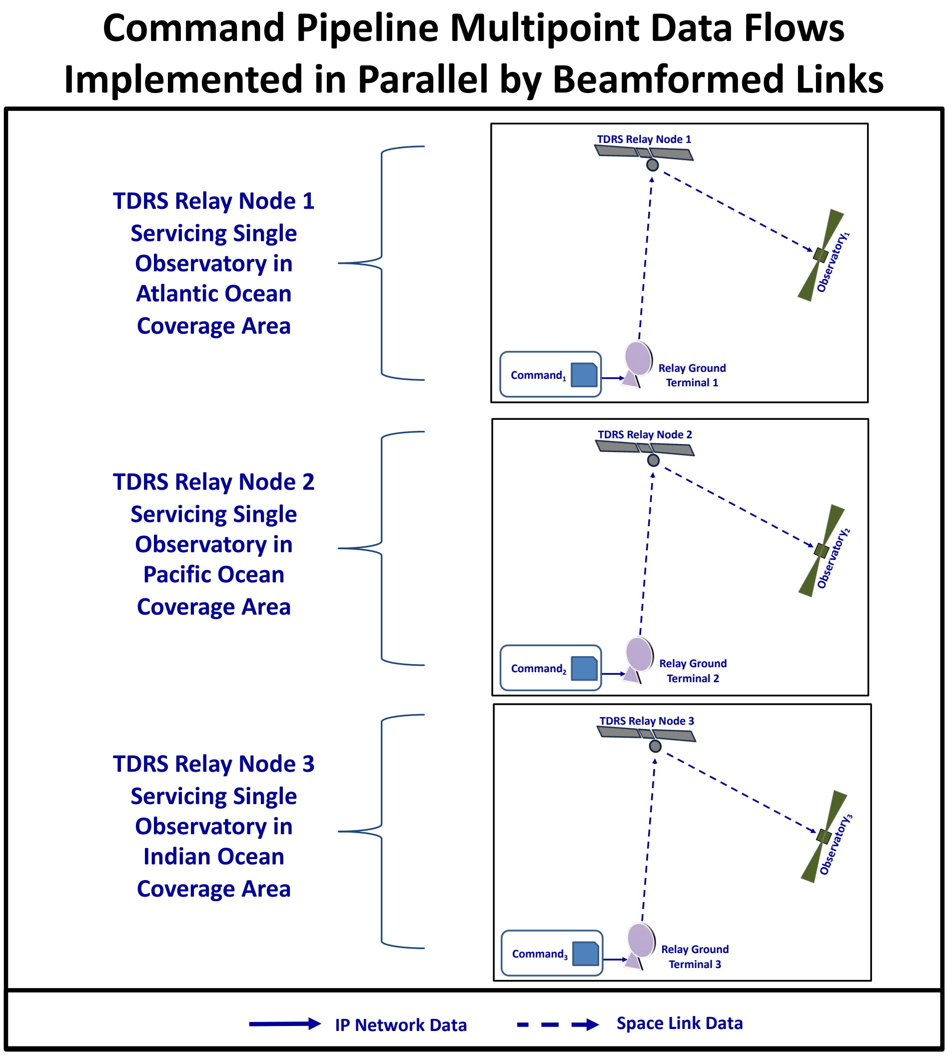


Fig. 19: Physical view of command pipeline multipoint data flows implemented in parallel by beamformed links.

As a result of the parallel servicing enabled by the new orbital phasing, the service start wait time, $SSWT_i$, for the three observatories is equal to the minimum service start lead time, *Min_Lead*. The time to complete, TTC, data flows in the scenario instance equals the minimum service start lead time, *Min_Lead* plus the minimum service execution period, *Min_SE.* The revised models for *SSWT* and *TTC* are given by Eq. 17 and Eq. 18 respectively.

$$SSWT_i = \text{Min\_Lead} \tag{17}$$

$$TTC = SSWT_i + \text{Min\_SE} \tag{18}$$

The revised models for *SSWT* and *TTC* impact the outcomes for the network utilization, *NU*, (Eq. 12) and effective data rate, *EDR*, (Eq. 14). No other system suitability metrics are impacted by the revised scenario implementation.

Table 15 provides a comparison of results for the point of departure (original) and revised command pipeline implementations.

Table 15: Comparison of results for the original and revised command pipeline scenario implementations.

Suitability Metric Category	Metric	Command Pipeline Multipoint Data Flows Implemented Sequentially by Beamformed Links	Revision 1: Command Pipeline Multipoint Data Flows Implemented in Parallel by Beamformed Links	Units
Timeliness Metrics	$SSWT_i$	$SSWT_1$ = 360 $SSWT_2$ = 450 $SSWT_3$ = 540	360	Seconds
	TTC	600	420	Seconds
Capacity Metrics	*NSI*	3	3	Number
	NU	0.46 %	0.12 %	Percent
Throughput Metrics	*DV*	2,400	3,400	Bits
	EDR	3.2	5.7	Bits per Second
	LTDV	12,000	12,000	Bits per Day

The revised observatory orbit phasing allows parallel servicing of each observatory in the scenario. The service start time for Observatory_3 is reduced by 33%. The time to complete the

scenario instance is reduced by 180 seconds, providing a 30% margin for satisfaction of the 600 second maximum latency requirement. The network utilization is improved by 74% and the effective data rate is improved by 44%.

Revision 2: Command and Notification Pipeline Implemented in Parallel by Apportioned Broadcast Link

The full bandwidth of the TDRS broadcast link was assumed to be allocated for implementation of the notification pipeline in Step 2c of the data flow design method. Broadcasting a standard notification message to all users simultaneously provides the most timely and efficient use of the broadcast link bandwidth. However, this approach requires that all users agree to standardize on the notification message and implement onboard autonomy to respond appropriately to the notification message contents. The second revision allocates both command and notification traffic to the broadcast link. The advantage of this revised implementation is that user missions may improve their timeliness outcomes by avoiding the wait times associated with discrete beamformed link servicing without the need for additional onboard autonomy or achieving consensus within the larger science community on the standardized notification message contents. The disadvantage is that apportioning broadcast service bandwidth to disseminate a unique command message to a single user is inefficient, since the unique command is of no value to all other users of the broadcast link. A physical view of the revised scenario is provided in Fig. 20.

Command and Notification Pipeline Multipoint Data Flows Implemented in Parallel by Apportioned Broadcast Link

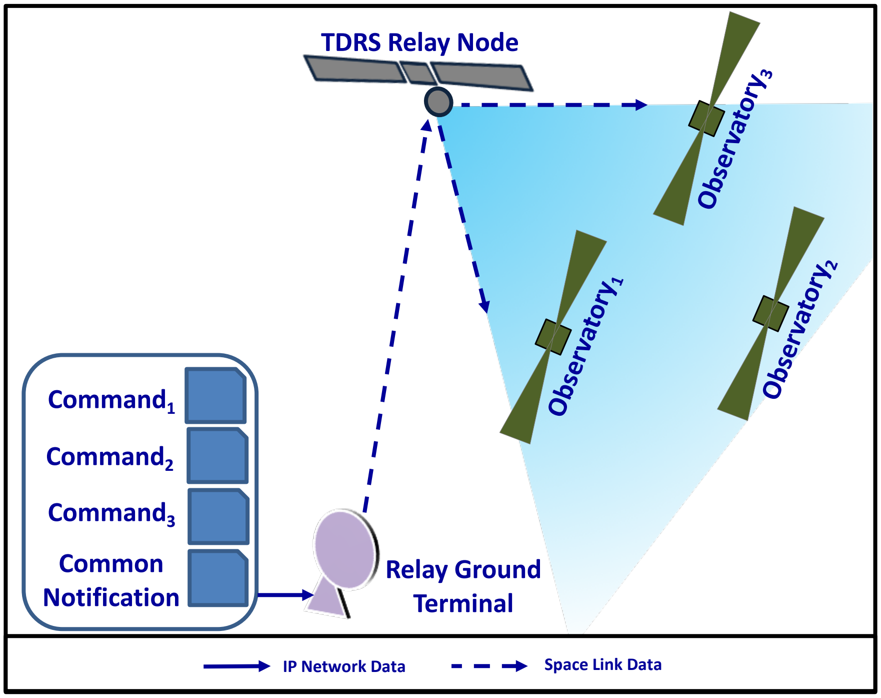


Fig. 20: Physical view of command and notification multipoint data flows implemented in parallel by a broadcast link.

The TDRS broadcast link bandwidth is apportioned into independent channels using a time-division multiplexing technique. The channels are serviced for a fixed period in a repeating sequence until the command or notification data assigned to each channel is delivered to the space-based observatory. In the apportionment scheme proposed by Heckler, four channels are dedicated to ground-to-space commands or notifications for an effective simultaneous channel servicing rate, *SSR*, of 20 bits per second per channel [36].

Complex queueing behaviors may occur in scenarios that require usage of more than four channels, where the data volumes of each command or notification message are variable, and when command or notification traffic for other scenarios must be serviced. These factors are not

considered in this design study for simplicity but are important considerations for the design of a future TDRS broadcast service. In addition, the small latency contributions at each user mission operations control center for building the unique commands and from the network notification pipeline processing are ignored. Consequently, the scenario results presented subsequently represent the best possible outcomes under the implementation assumptions of the apportioned broadcast service.

The revised scenario allocates user commands to three channels and a common notification message to the fourth channel. This approach allows unique commands to be sent to the three space-based observatories in the baseline scenario and allows any future space-based observatories to participate in the scenario provided they adopt the standard common notification message.

Under the stated assumptions, the scenario service start time for all space-based observatories occurs simultaneously and immediately following announcement of the neutron star merger transient event to the terrestrial transient science messaging application. The total scenario data volume, *DV*, is calculated by Eq. 13 and includes the data volume of each unique command plus the common notification message. Per the scenario specification, the data volume of each command sent by each of the command pipelines to their respective observatories and the data volume of the notification, DV_i, are each equal to 800 bits. As a result, each channel of the broadcast service completes its ground-to-space data transport servicing task at the same time. The revised model for the scenario time to complete, *TTC*, is given by Eq. 19.

$$TTC = \frac{DV_i}{SSR} \tag{19}$$

A single broadcast link network service instance, *NSI*, is required to transport the scenario data, as in the original dedicated broadcast service implementation.

The long-term network utilization the revised scenario, *NU*, is provided by Eq. 12 using the revised model for *TTC* (Eq. 19) and a value of zero for the minimum service start lead time, *Min_Lead*, due to the continuous link provided by the broadcast service implementation.

The effective scenario servicing data rate, *EDR*, is calculated by Eq. 14 using the total command and notification data volume delivered for the revised a scenario instance, *DV*, and the revised model for scenario time to complete, *TTC* (Eq. 19).

The model for the long-term scenario data volume throughput, *LTDV*, is provided by Eq. 15 using the total command and notification data volume, *DV*, delivered for the revised scenario instance.

Table 16 provides a comparison of results for the point of departure (original) notification pipeline broadcast link implementation and the revised broadcast link implementation with bandwidth apportioned for command and notification pipelines.

Table 16: Comparison of results for the original and revised broadcast link scenario implementations.

Suitability Metric Category	Metric	Notification Pipeline Multipoint Data Flows Implemented in Parallel by Broadcast Link[17]	Revision 2: Command and Notification Pipeline Multipoint Data Flows Implemented in Parallel by Apportioned Broadcast Link[17]	Units
Timeliness Metrics	$SSWT_i$	0	0	Seconds
	TTC	0.78	40	Seconds
Capacity Metrics	*NSI*	1	1	Number
	NU	0.002 %	0.08 %	Percent
Throughput Metrics	*DV*	800	3,400	Bits
	EDR	1,024	80	Bits per Second
	LTDV	4,000	16,000	Bits per Day

The revised implementation of command and notification pipelines by an apportioned broadcast link allows users to retain the traditional architecture for commanding space-based observatories at the expense of less efficient outcomes on nearly all system suitability metrics as compared to the original broadcast implementation of a notification pipeline. The revised *TTC* of 40 seconds is more than 50 times greater than the dedicated broadcast service implementation. However, the revised *TTC* is still 15 times faster than the maximum allowable scenario *TTC* requirement of 600 seconds. The increases to the scenario instance data volume (*DV*), long-term data volume (*LTDV*) and network utilization (*NU*) are due to the unique commands for each observatory in addition of the common notification data in the revised implementation. Inclusion of common notification data in the revised scenario provides an infusion pathway for future

observatories to participate in the scenario without further impact to the system suitability metrics. Finally, the effective scenario servicing data rate, *EDR*, of the TDRS broadcast service was reduced from the maximum achievable rate of 1024 bits per second to a proposed four channel apportionment, resulting in an effective scenario servicing data rate of 80 bits per second.

Discussion of Results

Terrestrial science messaging applications enable dynamic and coordinated observations of transient scientific phenomena among diverse and spatially distributed observatories. Information is exchanged among observatories using publish and subscribe messaging relationships, which define multipoint data flow patterns within a communications network. Science messaging applications use categorical topics and data message contents specific to the scientific objectives of the community they serve. Some ground-based observatories have implemented machine-to-machine interfaces with science messaging applications to automate timely follow-up observations in response to information received from the science messaging application. Terrestrial science messaging applications can implement timely multipoint data flows using the addressing, routing, multicast and differentiated datagram packet transport capabilities of Internet Protocol-based networks.

Direct extension of science messaging applications to space-based observatories is constrained by current space data network protocols and manual procedures to request, provision and execute space data transport services. Traditional space data network design and operations are oriented towards servicing pre-planned and uncorrelated data transport needs for individual users. By contrast, transient science observation scenarios impose randomly

occurring and correlated user demand for time-sensitive multipoint data flows on space data networks. Such scenarios may impose network access precedence or sequencing requirements among the observatories. Multi-tiered and conditionally branching scenarios based on information in the user data stream are also possible.

This research has made several advances to facilitate the design and evaluation of multipoint data flows used in transient science observation scenarios. First, the command pipeline and notification pipeline functional reference architectures provide alternative approaches for implementing timely multipoint ground-to-space data flows in the context of current space data network systems, protocols and procedures. As a compliment to existing timely space-to-ground communications services, the reference architectures enable space-based follow-up observations in end-to-end transient science observation scenarios. Second, a multipoint data flow design and evaluation method based in queueing theory was defined and applied to transient science observation scenarios. Third, a design study for a notional multi-observatory transient science scenario was performed to demonstrate how the impacts of alternative command and notification pipeline implementations can be quantified and evaluated using the multipoint data flow design method and system suitability metrics. A summary of the notional design study results is provided in Table 17.

Table 17: Summary of results for alternative implementations of the transient science observation scenario.

Suitability Metric Category	Metric	Command Pipeline Multipoint Data Flows Implemented Sequentially by Beamformed Links	Revision 1: Command Pipeline Multipoint Data Flows Implemented in Parallel by Beamformed Links	Notification Pipeline Multipoint Data Flows Implemented in Parallel by Broadcast Link[17]	Revision 2: Command and Notification Pipeline Multipoint Data Flows Implemented in Parallel by Apportioned Broadcast Link[17]	Units
Timeliness Metrics	$SSWT_i$	$SSWT_1$ = 360 $SSWT_2$ = 450 $SSWT_3$ = 540	360	0	0	Seconds
	TTC	600	420	0.78	40	Seconds
Capacity Metrics	*NSI*	3	3	1	1	Number
	NU	0.46	0.12	0.002	0.08	Percent
Throughput Metrics	*DV*	2,400	2,400	800	3,400	Bits
	EDR	3.2	5.7	1,024	80	Bits per Second
	LTDV	12,000	12,000	4,000	16,000	Bits per Day

Command Pipeline Discussion

The command pipeline reference architecture preserves the current functional allocations between the user and space data network service provider system domains. As a result, command pipelines for simple transient science scenarios can readily be implemented using space data networks that provide a real-time service access request and provisioning interface, including the TDRS space data network and emerging commercial direct-to-Earth network service providers. In collaboration with this research, the Swift mission implemented

the first known command pipeline for science applications in January 2020. The Swift command pipeline issues a command to the observatory to archive raw instrument data in response to a message received through the GCN science messaging application indicating that ground-based observatories have detected a gravitational wave transient event. The Swift command pipeline has increased the detection rate of gamma-ray bursts associated with ground-observed gravitational wave events by 400% [53]. More sophisticated scenarios involving multiple space-based observatories may require new queue management mechanisms to coordinate and enforce network access precedence or sequencing rules required by the scenario. In addition, implementing command pipeline data flows using discrete beamformed links to each space-based observatory has limitations as the number of space-based observatories involved in transient science scenarios increases. Such limitations are illustrated in the Table 17 results. The outcomes on the system suitability metrics for command pipelines implemented by beamformed links, whether sequentially by a single TDRS node or in parallel by multiple TDRS nodes, have significantly worse outcomes on nearly every system suitability metric when compared to the command pipeline implemented by an apportioned TDRS broadcast link. A TDRS beamformed link can provide 15,000 times greater link data rate than the apportioned broadcast link data rate (300 kilobits per second versus 20 bits per second, respectively). However, very little data volume is required to encode an observatory command (approximately 800 bits). The high data rates of beamformed links provide little benefit when transporting small data volume commands. The latencies induced by the long lead time to request, provision, and execute data transport services for beamformed links dominate the latencies induced by the transport of commands by a very low but continuous data rate provided by an apportioned broadcast link. Consequently, although

the specific suitability metric outcomes for different multi-observatory scenarios using the TDRS space data network will vary, the relative suitability ranking of the three command pipeline implementations explored in the design study will hold true for typical observatory command data volumes. In summary, a scenario involving command pipelines used by any number of space-based observatories will execute most efficiently using an apportioned broadcast implementation, and scenarios where the observatories are spatially distributed to allow parallel servicing with beamformed links will execute more efficiently than those with spatial distributions that necessitate servicing multiple observatories sequentially with a single TDRS node.[18]

A disadvantage to using an apportioned broadcast link for dissemination of unique command data is the opportunity cost of the link bandwidth, since only a single observatory can make use of the command data that is disseminated to all space-based observatories. As the number of unique commands disseminated by the broadcast link increases, the queue servicing time will also increase, eventually reaching a point of impracticality. Notification pipelines provide a means to address the inefficient use of broadcast link bandwidth apportioned for unique commands by disseminating common standardized notification messages that can be used by multiple subscribed space-based observatories.

Notification Pipeline Discussion

The notification pipeline reference architecture introduces several changes to traditional user and space data network functions, and more closely approximates the publish-subscribe

[18] For scenarios involving observatories in various orbits, simulations can be used to characterize the varying number of observatories that must be serviced sequentially or can be serviced in parallel over a defined observation campaign period, as described in Step 2a of the multipoint data flow design method.

behaviors found in terrestrial science messaging applications. First, the space data network subscribes to messages published to a terrestrial science messaging application. The notification pipeline processes terrestrial science messages into secure standard common space data notification message. Then, the notification pipeline implements, by proxy, user preferences about the notification messages based on science topic, observed event attributes, or other characteristics. Next, the notification pipeline performs traditional network access activities to disseminate the notification message to space-based observatories. In this way, the notification pipeline simplifies user access and provides timely multipoint data flows for transient science observation scenarios. In addition, the notification pipeline could implement complex network access precedence or sequencing requirements established by the users involved in the transient science scenario using a rule-based data routing function. However, notification pipelines present new challenges for security, integration testing and validation. Robust fault detection and isolation controls are needed to mitigate risks of unanticipated emergent behaviors arising from autonomous interactions among the user and network nodes. The notification pipeline requires consensus among the science community of interest on standard notification message protocols and contents. Additionally, the notification pipeline requires allocation of follow-up decision making in response to notification messages received by the space-based observatory. The use of common standardized notification messages and a dedicated broadcast link implementation enables the most efficient servicing of all subscribed space-based observatories, as illustrated by the results in Table 17. However, a practical broadcast link would need to apportion bandwidth to allow servicing of multiple simultaneous transient science scenarios or to concurrently transport other types of time-sensitive multipoint notification traffic.

Nevertheless, the use of a broadcast link to disseminate common standardized notifications for use by multiple space-based observatories will always be more efficient than apportioning the broadcast bandwidth to unique commands for a single observatory. In principle, notifications could be disseminated by beamformed links, although this approach would lessen the parallel servicing advantages and scalability of a broadcast link.

Demonstration & Practical Considerations

In the notional design study, it was assumed that TDRS data transport services were available without blocking from pre-scheduled users or contention for access with other transient science observation scenarios. This assumption allows distinctions between the alternative beamformed link and broadcast link servicing approaches to be clearly seen in the Table 17 results. Presently all 10 operational TDRS nodes can support one service instance using the forward antenna array. Each of the six later generation TDRS nodes could support up to two simultaneous forward service instances with ground segment upgrades. An analysis of TDRS forward service utilization presented in Chapter 3 and in Roberts et. al. suggests that it could be possible to implement a policy that allocates pre-scheduled user service periods across the TDRS constellation to eliminate or minimize blocking for transient science observation scenarios [46]. As a result, Near Space Network planners likely have operational and engineering development options available to implement transient science scenario data flows using beamformed and broadcast links without adverse impacts to pre-scheduled users.

Presently, the Near Space Network has no official requirements or memoranda documenting transient science space system user needs. There is no stakeholder consensus to implement a notification pipeline nor an operational broadcast service. Additionally, there is no

policy to allocate pre-scheduled service periods across the TDRS constellation to minimize blocking for transient science users. As a result, current users seeking to implement transient science observation scenarios are limited to beamformed link implementations of command pipelines and must contend with blocking due to pre-scheduled services by other TDRS users. The expected latency introduced due to blocking by pre-scheduled TDRS users was explored and quantified in Chapter 3 and in Roberts et. al. [46]. The impacts of contention with pre-scheduled TDRS users should be incorporated into the queueing behavioral models as necessary to reflect the queue servicing factors of the TDRS constellation when evaluating any proposed multi-observatory transient science scenario.

Swift's results from its implementation of the first command pipeline demonstrate the value of extending publish-subscribe behaviors to space data networks and serve as an exemplary pathfinder for the further development of transient science observation scenarios. Based on the TDRS servicing models developed in Section 3, the maximum command pipeline servicing capacity per TDRS node is estimated to be approximately 40 space-based observatories if implemented by sequential beamformed links and approximately 325 space-based observatories if using the apportioned broadcast link. Considering the blocking probability and duration of blocking periods for beamformed links, each TDRS node has a maximum servicing capacity of approximately 10 space-based observatories. This implies that the current state of the TDRS space data network has sufficient capacity to support the operations of additional multi-observatory transient science scenarios using command pipelines implemented by sequentially beamformed links provided their scenario time to complete requirements are on the order of 10 to 90 minutes. To satisfy scenarios with timeliness requirements of less than 10 minutes or for

scenarios involving larger numbers of space-based observatories, an apportioned broadcast link implementation of command pipelines is needed. As the value of multi-observatory transient science concepts becomes more readily apparent, the apportioned broadcast link could transition from transporting command pipeline data to notification pipeline data over time to maximize efficiency.

Conclusion

The command and notification pipeline reference architectures developed by this research reflect the particularities of current space data network protocols and procedures. Such protocols and procedures are rooted in an intermittently connected and pre-reserved circuit-switched service delivery paradigm for individual users within a single service provider network. Command and notification pipelines orchestrate the “out-of-band” service access processes used by space data networks to achieve multipoint space data flows. By contrast, terrestrial Internet Protocol networks use a continuous connectivity and packet-switched service delivery paradigm to route data between any source and destination on the network using federations of diverse network service providers.

The Near Space Network is presently leading commercialization and interoperability efforts for direct-to-Earth and space relay network services to provide users with more diverse multi-network service provider implementation options for their mission data flows. NASA and other international space agencies have endorsed Bundle Protocol to transition space data networks to a packet-switched service delivery paradigm and overcome the unique challenges of operating wireless communications networks in the space domain. A concise summary of end-to-end data flow use cases and current space-terrestrial interoperability approaches for Internet

Protocol and Bundle Protocol networks can be found in [50] and [54]. Future work is needed to make seamless interoperability between space and terrestrial communications networks a reality.

Chapter 5

Chapter 5 addresses Research Question #3:

How do commercial service provider network solutions compare to TDRS for transient science space systems?

Introduction

The National Academy of Sciences has identified time-domain multi-messenger astrophysics – the coordinated observation of transient scientific sources with complimentary ground and space-based observatories – as the highest-priority sustaining activity for the 2020’s [55]. Advances in space mission planning methods and network infrastructure to enable time-sensitive observations are necessary to realize the full potential of multi-messenger astrophysics [1].

The NASA Near Space Network (NSN) was chartered in October 2020 to provide a single point of contact for space data network service planning and operations for user missions venturing up to two million kilometers from the Earth [56]. The NSN seeks to leverage commercial capabilities to increase the efficiency and robustness of NASA-owned infrastructure [57]. Although the NSN will continue to rely on services from some government owned assets, including the Tracking and Data Relay Satellite (TDRS) space relay constellation, NSN has been chartered to take a “commercial first” approach for meeting the requirements of future mission users.

Over the course of several decades, NASA has developed and sustained robust modeling and simulation capabilities for its government-owned space data network assets [21]. These capabilities have been used to design and evaluate network solutions for NASA's mission users [58]. However, detailed technical models of commercial network assets are often proprietary. Furthermore, the commercial market for direct-to-Earth and space relay network services is rapidly evolving. The NSN has prototyped a new modeling and simulation tool to address these challenges. The tool presently allows network planners to rapidly evaluate potential solutions to satisfy mission user needs for commercial direct-to-Earth and commercial space relay network topologies. It relies on a set of moderate-fidelity regression models derived from orbit and communications link simulations of commercial network assets. The modeling and simulation parameters for the commercial network assets are based solely on publicly available data sources, including trade journal articles, conference proceedings, public filings with the Federal Communications Commission and data from company websites. The goal of the tool is to allow mission and network planners to rapidly explore and identify potential commercial network solutions for more detailed, and likely proprietary, investigations. This approach maximizes the commercial network solution trade space while protecting the proprietary information of commercial service providers. In addition, the tool helps ensure replicability of commercial network solution trade space analyses despite the knowledge firewalls placed between network planners working on mutually competitive science mission proposal teams.

State of the Practice

The TDRS Demand Access System (DAS) provides timely space-to-ground communications services, allowing space-based astrophysical observatories such as Swift and Fermi to alert

ground-based observatories of transient scientific events within seconds of detection using terrestrial science messaging applications [32] [16] [41]. Rapid space-to-ground communications service response times are achieved through use of dedicated (i.e., continuously reserved and allocated) ground-based communications signal beamformers for each DAS user. However, the ground-to-space (i.e., forward) communications signal beamformers are located onboard the TDRS spacecraft. Presently, the onboard beamformers have the capacity to support a single ground-to-space service instance at a time.[19] TDRS capacity constraints and resource contention among multiple users currently limit the achievable response time for ground-to-space communications services. However, ground-to-space service capacity constraints could be alleviated with planned upgrades to the TDRS ground system [59].[20] In 2016, Heckler et al. demonstrated that a continuous ground-to-space TDRS broadcast service with a usable bandwidth of 1,024 bits per second is feasible with additional TDRS ground enhancements [36].[21]

Mission operators need timely ground-to-space communications services to send commands and other data to space-based observatories to perform follow-up observations of

[19] In addition to the electronically steered arrays, each space-based TDRS node has two high gain mechanically steered antennas capable of bidirectional communications. The TDRS high gain antennas are not suitable for routine time sensitive commanding applications due to their higher utilization and slower service setup latency. As a result, use of the mechanically steered antennas is not considered in this research.

[20] The second and third generation space-based TDRS nodes were designed with the capability to split the forward antenna array into two independent beams. With ground system upgrades, each second and third generation TDRS node could support two simultaneous ground-to-space service instances.

[21] Splitting the forward array reduces the available power for each link by approximately three decibels, translating to a reduction in the maximum achievable communications link data rate. The relatively low broadcast data rate results from the power losses due to splitting the TDRS forward beam and transmitting into a wide volume covering all low-Earth-orbit users simultaneously. The bandwidth of the broadcast service could be allocated for individual user commands or for common notification messages that aid in autonomous mission operations. Ground enhancements to integrate the data streams onto the broadcast service are required. In addition, users are required to maintain continuous connectivity with the broadcast service to eliminate service setup and signal acquisition latencies. Users must also manage connectivity handovers across TDRS nodes during their orbit. Pre-scheduled services using the second independent TDRS beam would be performed as they are presently but have lower available signal power due to the splitting of the forward array.

transient scientific events. In 2020, Swift implemented the first known autonomous command pipeline for a space-based observatory [53]. Following receipt of a transient science event notification from a terrestrial Internet Protocol-based science messaging application, the Swift command pipeline autonomously generates a follow-up observatory command, schedules a TDRS ground-to-space communications service instance and sends the command to the Swift observatory in as little as 14 minutes [53]. However, since the scientific value of a follow-up observation decreases with time, minimizing ground-to-space data delivery timeliness is essential. Two non-NASA technology pathfinder missions identified in the literature are planned to evaluate the performance of commercial space data networks for transient astrophysics applications. The AGU Remote Innovative CubeSat Alert (ARICA) satellite was launched in November 2021 as part of the Japan Aerospace Exploration Agency (JAXA) Innovative Satellite Technology Demonstration Program [60] [61]. The University of Melbourne is developing the Space Industry Responsive Intelligent Thermal satellite (SpIRIT) for launch in 2022 [62]. ARICA and SpIRIT will fly both Iridium and Globalstar mobile terminals to evaluate the communications response time of each network.

Iridium and Globalstar differ in significant ways from the TDRS space data network. TDRS was designed to provide global coverage to a limited population of government-owned space-based mission users according to a pre-determined service schedule. By contrast, Iridium and Globalstar are personal satellite communications networks designed to provide on-demand service access to large populations of commercial Earth-based users. Several efforts to model and simulate the expected coverage of commercial personal satellite communications networks for space-based users have been undertaken in the literature [63] [64] [65]. However, results

from these efforts vary widely in part due to differing assumptions about the design characteristics of user and provider nodes (i.e. antenna, signal and link parameters), the performance of proprietary service initiation and mobility management protocols, and the effects of the relative orbital dynamics (e.g. range, velocity, doppler, etc.) between user and provider nodes over time. The literature also provides data from ground and flight testing of Globalstar and Iridium networks for space-based users [66] [67].

Research Objectives

The research objectives of Chapter 5 are 1) to extend the descriptive model of a transient science space system developed in Chapter 3 and Chapter 4 to incorporate commercial service providers, 2) to develop quantitative models estimating system outcomes with the Globalstar and Iridium commercial service providers and 3) to compare the relative suitability of the as-built TDRS network with the commercial Globalstar and Iridium network solution options for a hypothetical design study of a transient science follow-up observation scenario.

System Model Development

A hypothetical design study for a transient science observation scenario is extended using a multipoint data flow design and evaluation method.[22] The four major steps of the method are illustrated in Fig. 21.

[22] The hypothetical design study and scenario details are defined in Chapter 4, Section 3.

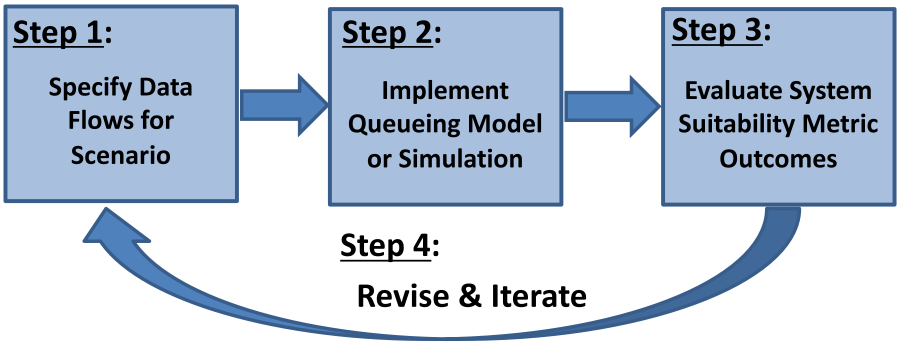


Fig. 21: Data flow design method for transient science space system design studies (Chapter 4).

In Step 1, abstract system model views of a transient science space system incorporating commercial service providers are developed using Systems Modeling Language (SysML). In Step 2, quantitative models to estimate suitability metric outcomes using Globalstar and Iridium commercial networks for the scenario data flows are derived from the literature and computational simulations. In Step 3, results from the commercial network service provider models are compared with those of the TDRS network.[23] The optional Step 4 is not demonstrated here.

Step 1: Specify the data flow characteristics of the scenario.

The NSN serves as mission users' single point of contact for network planning and operations across government-owned and commercial service provider space data networks. The NSN corresponds to a level in the abstract system hierarchy that sits above the functions of any individual space data network. Accordingly, the NSN's service management functions go beyond those of a single space data network since they must broker and orchestrate the data transport services executed by others.

---

[23] Quantitative TDRS models and results are presented in Chapter 3 and Chapter 4.

In this hypothetical design study, the NSN serves as the network service provider. Mission and NSN analysts must evaluate the suitability of government owned (TDRS) and commercial space data network options to implement timely, randomly occurring, command data flows for a transient science scenario.

System Structure

The command pipeline functional reference architecture preserves the traditional allocation of user and network service provider functions, making it suitable for implementation with commercial space data networks.[24] Based on a review of the literature, the command pipeline functional reference architecture is also consistent with the interfaces and constraints of both the Globalstar and Iridium commercial networks [66] [67].[25]

An abstract structural model of a transient science space system with a command pipeline is extended to include one or more commercial service providers in Fig. 22.

[24] The command pipeline functional reference architecture was developed in Section 1 of Chapter 4.

[25] Although implementation details vary, Globalstar and Iridium support an Internet Protocol-based application interface with terrestrial users (mission operations centers). Both commercial networks can provide timely and secure bi-directional data flows with a mobile (space-based) terminal. Qualified space-based user terminals exist for each commercial network and fit within the typical design accommodation budgets of small satellites. The literature also addresses regulatory approval and licensing processes for government users of commercial personal satellite communications networks with national and international spectrum authorities. The low data volume required to encode commands for a typical spacecraft (800 bits in the hypothetical design study) are suitable for the short burst text messaging services offered by each commercial provider. In addition, the space-to-ground data flow provides an option for ground mission operators to verify receipt of commands sent to the spacecraft. The suitability of Globalstar and Iridium space-to-ground data flows for command verification and transport of science mission data are not considered in the scope of this hypothetical design study.

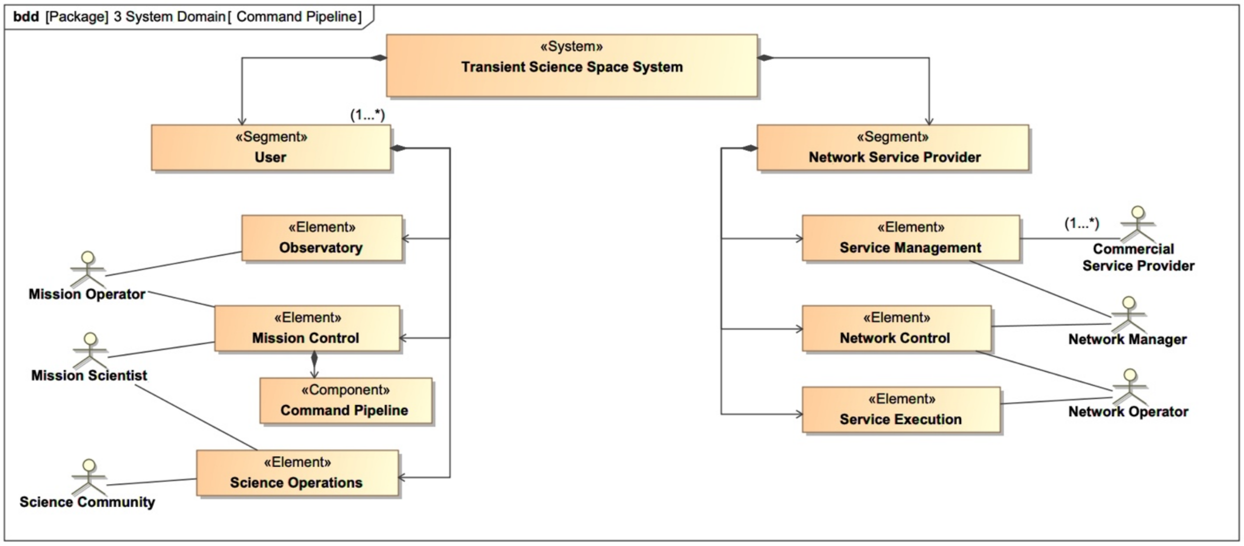


Fig. 22: Structural model of a transient science space system with commercial service providers.

The association between the network service provider's service management element and the external commercial service provider actor reflects the NSN's role as a broker of commercial services. The network control and service execution functions and their associated actors pertain to government-owned network assets that provide services to NSN, such as the TDRS space data network. As in Chapter 4, the command pipeline is illustrated as an extension to the user mission control element.

Activity Flow and System Allocation

Command pipelines accelerate follow-up observations of transient scientific events for space-based observatories. Command pipelines generate and send commands from the users' terrestrial mission operations centers to their respective space-based observatories in response to notifications of transient scientific events received from a terrestrial Internet Protocol-based science messaging application.

An abstract behavioral model of the functional allocation of activities for a command pipeline data flow involving a commercial service provider is provided in Fig. 23.

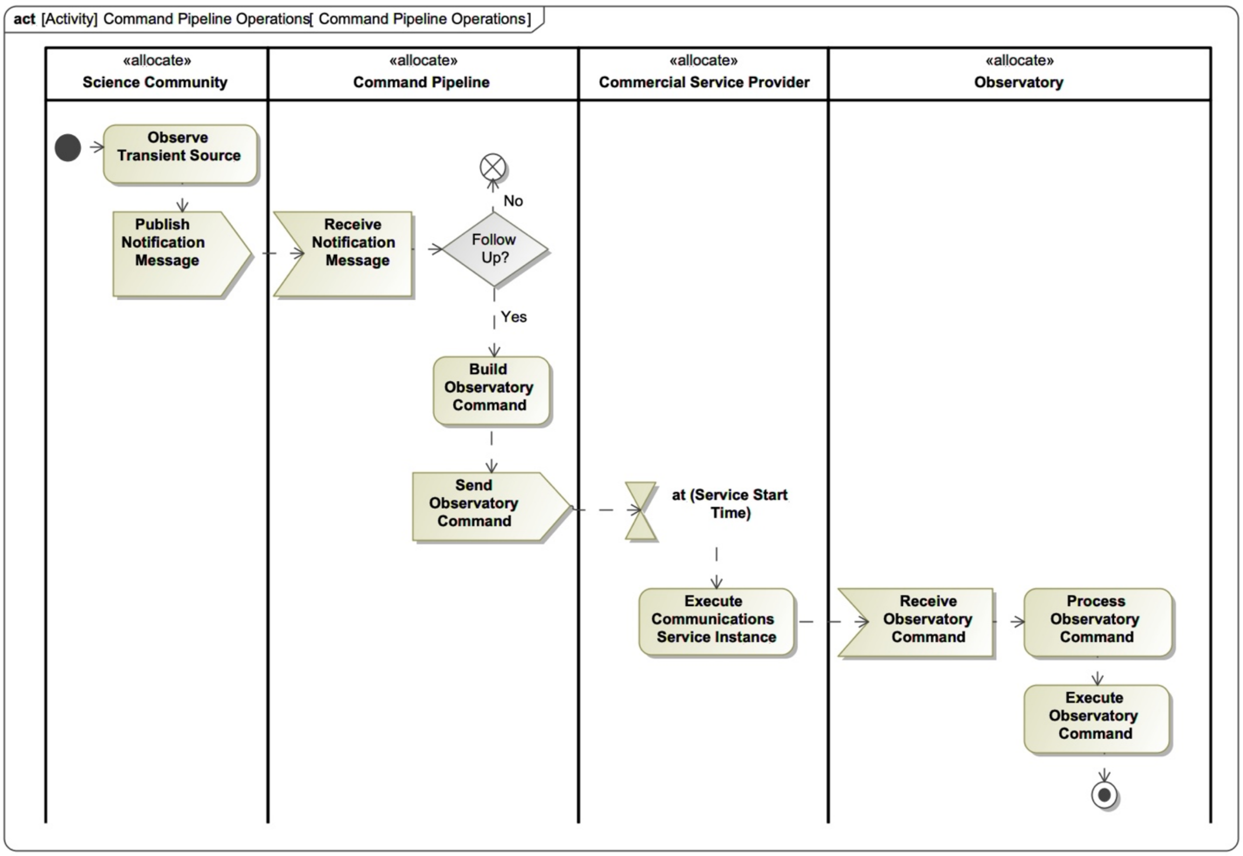


Fig. 23: Activity flow of a command pipeline using a commercial space data network service provider.

The literature provides data to estimate the queue wait time and data flow service instance execution time for the activities allocated to the commercial service provider.

End-to-End Operational View

A hypothetical transient science observation scenario was defined in Section 3 of Chapter 4. The end-to-end operational view of the scenario has been extended to illustrate the three space data network provider options in Fig. 24.

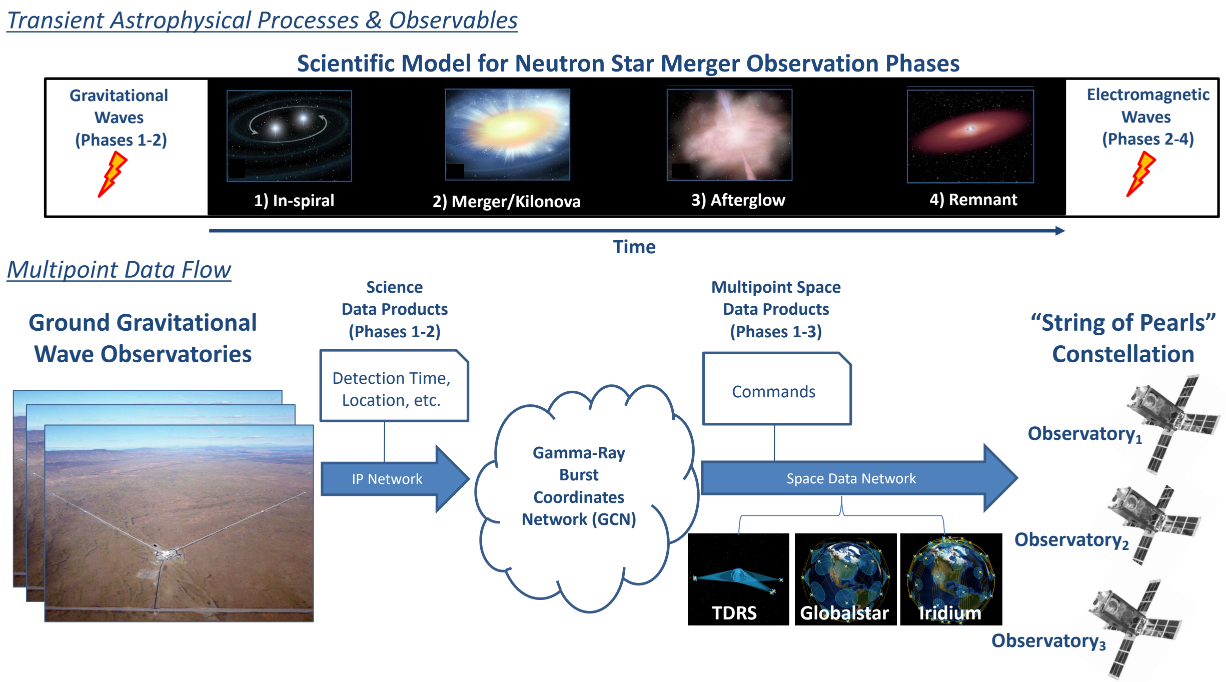


Fig. 24: Operational view of a transient science observation scenario illustrating three space data network options.

There are no further extensions to the abstract system models and data flow characteristics of the scenario defined in Section 3 of Chapter 4.

Step 2: Implement system queueing behavioral model or simulation.

Transient science observation scenarios pose several new challenges to current space mission design practices. Queueing theory provides a quantitative foundation for assessing the suitability of alternative network implementations of the required data flows for a transient science scenario. Chapter 4 provides a systematic multipoint data flow design and evaluation method along with a set of system suitability metrics to address the system planning challenges posed by transient science observation scenarios. Also in Chapter 4, results are presented from a hypothetical transient science space system design study using the TDRS space data network. This section presents quantitative data from the literature and new modeling results to estimate

system outcomes using the Globalstar and Iridium commercial networks. In the literature, Voss et. al. provide a detailed discussion on the use of Globalstar network services for space missions, including ground and flight experimental designs and results [66]. Riot et. al. provide a similarly detailed discussion and experimental results for Iridium [67]. These references provide foundational information about each commercial service provider's system architecture, operational requirements and constraints, and data flow testing outcomes. The subsequent section interprets and applies information from these sources in the context of a hypothetical multi-observatory transient science scenario.

Queueing Processes in Globalstar and Iridium Commercial Networks

As developed in Chapter 4, the queue arrival process for a transient science scenario can be characterized by two factors: the mean occurrence rate of a transient scientific event (*Event_Rate*, specified at 150 occurrences per month) and the number of associated space data flows it initiates (*NSI*, specified at 3 network service instances per scenario instance). The number of associated space data flows (*NSI*) initiated by an occurrence of a transient scientific event depends on both the scientific observation strategy defined by the scenario design and implementation details of the data flows.[26]

In the hypothetical design study, the scenario defines ground-to-space data flows between each of the three ground-based mission operations centers and their associated space-based observatories. A representative command data volume of 800 bits per ground-to-space data flow is specified in Chapter 4. The small command data volume is suitable for encoding and

[26]For example, data flow implementation options explored in Chapter 4 include use of common notifications versus unique commands and use of point-to-point spot beams versus multi-point broadcast links.

transport using the Short Messaging Service (SMS) and Short Burst Data (SBD) services offered respectively by Globalstar and Iridium [66] [67]. The timeliness requirement for the scenario is specified as follows: “The command data shall be delivered as soon as possible to the space-based observatories and must be delivered no later than 10 minutes following initiation of service access processes for space data network services.”

To implement end-to-end data flows, the Iridium and Globalstar commercial service providers rely on low-Earth-orbiting space nodes, space-ground gateway stations and Internet Protocol-based terrestrial networks. Wireless mobile user terminals connect to the networks through non-tracking spot-beams from the space-based provider nodes. The commercial networks use proprietary signals and protocols for user registration, authentication, mobility management and other aspects of service access and execution. The commercial networks provide an application interface to Internet Protocol-based networks through which user ground control centers may send command data messages through the network to the user’s wireless mobile terminal onboard the space-based observatory. The commercial networks encode, store, forward and, if necessary, retransmit user command data received through the application interface using non-standardized procedures and protocols. The commercial networks are designed to handle randomly initiated user data flows and have the capacity to support large numbers of simultaneous user service instances. Accordingly, the commercial networks perform independent and parallel queue servicing of the user command data transport demands of the transient science scenario.

System Suitability Metrics

In Chapter 4, three technical suitability metric categories for evaluating alternative space data network solutions are defined.[27] In this section financial cost is added as a fourth suitability metric category.

The total financial cost of implementing a space data network solution can be estimated based on three component metrics. First, C_{Dev}, is the non-recurring engineering development and/or commercial procurement costs for the space and ground user-network interfaces (e.g., space-based user terminal and ground interface hardware and software). Second, $C_{I\&T}$, is the non-recurring cost for integration and testing to demonstrate compatibility and successful data flows across user and network systems. Third, $C_{OM\&S}$, is the recurring cost for operating, maintaining and sustaining data services (e.g., periodic service charges and costs for performing software patching and security updates).

The expanded set of system suitability metrics is provided below in Table 18.

[27] The three technical metric categories are timeliness, capacity and throughput.

Table 18: Space data network suitability metrics for transient science space system design studies.

Suitability Metric Category	System Suitability Metric Definition	Notation
Timeliness Metrics	Service start wait time for the i^{th} observatory in the scenario	$SSWT_i$
	Time to complete scenario instance data flows	TTC
Capacity Metrics	Number of network service instances per scenario instance	NSI
	Long-term space data network utilization imposed by the scenario	NU
Throughput Metrics	Scenario instance data volume throughput	DV
	Effective scenario servicing data rate	EDR
	Long-term scenario data volume throughput	$LTDV$
Financial Cost Metrics	Development and/or procurement cost of user-network interfaces	C_{Dev}
	Integration and testing cost of user-network interfaces	$C_{I\&T}$
	Operations, maintenance and sustainment cost of services	$C_{OM\&S}$

Timeliness Modeling

Network timeliness outcomes depend on intrinsic network latency factors and on scenario-dependent service gap periods.

Intrinsic Network Latency Factors

The minimum timeliness outcomes achievable by a network depend on intrinsic latencies associated with factors such as service initiation protocols, setup and routing delays and data transport rates. Generally, the intrinsic latency of a network, *Intrinsic_Latency*, is a random variable whose distribution can be characterized by sending representative messages between endpoints of the network when there is end-to-end connectivity (i.e., a service gap does not occur). For simplicity, the mean value for intrinsic latency is used to model the minimum timeliness outcome in this hypothetical design study. More sophisticated analyses may use Monte Carlo techniques to sample the distribution of minimum timeliness outcomes. This

approach is especially well-suited for end-to-end system modeling where the characteristics of the timeliness distribution may have significant impacts on overall system effectiveness outcomes.

For the Globalstar network, in 2016 Voss et al. reported results from ground testing of Globalstar SMS mobile-terminated data flows. An automated script initiated 244 representative command message data flows from an emulated user mission control center to a ground-based wireless mobile terminal. The reported mean time to transport the message was 28 seconds. As a result, *Intrinsic_Latency* of the Globalstar network equals 28 seconds.

For the Iridium network, in 2021 Riot et al. reported results of similar but less extensive ground testing (10 data flow trials) characterizing the inherent latency of Iridium SBD mobile-terminated data flows. Although the distribution of outcomes is not published, the range of observed message transport latencies were between 10 and 20 seconds. These results are consistent with mobile-terminated SBD data flow ground test results published in 2005 by McMahon and Rayburn [68]. The midpoint of the range of observed values (15 seconds) is used for the Iridium mean intrinsic network latency. As a result, *Intrinsic_Latency* of the Iridium network equals 15 seconds.

For the commercial networks, the contribution of the *Intrinsic_Latency* variable to the *TTC* timeliness suitability metric is expressed subsequently in Eq. 22 and Eq. 23.

Factors that contribute to the intrinsic latency of the TDRS network were identified, analyzed and quantified in Chapter 3 and Chapter 4. The intrinsic latency of the TDRS network is dominated by a fixed minimum service start lead time. TDRS scheduling ground rules specify the minimum service start lead time to be 10 minutes from acceptance of a user service access

request message. Ground testing with the engineering instance of the TDRS service management and network control systems demonstrated that user access request messages specifying as little as six-minute lead times are accepted and scheduled, whereas requests specifying lesser lead times are rejected. The experimentally determined six-minute service start lead time is used for the purpose of the hypothetical design study. Two much smaller variable wait time contributors, the user's service access request message build time and the network's access request processing time, are estimated to contribute an additional mean wait time of 5 seconds. As a result the intrinsic latency for the TDRS network to start servicing a single TDRS network user, or the first user in a serial queue, is six minutes and five seconds.

Scenario-Dependent Network Service Gap Latency Factors

Scenario-dependent network service gap periods may occur due to planned or unplanned user or network downtime, resource blocking from other users, or failure to close the free-space communications link due to orbit geometry and other user and network design factors. Each of these contributing factors can be represented as a random variable with a characteristic distribution derived from empirical data sources. For the hypothetical design study, it is assumed that user and network nodes have no service gaps due to planned or unplanned downtime. Service gap characteristics and their impact on timeliness outcomes were quantified in Chapter 3 and Chapter 4 for the TDRS network. The probability of TDRS service gaps due to blocking was found to be 16%. TDRS blocking periods were found to contribute a mean wait time of 2 minutes and 58 seconds. Orbital and communications link simulation results in Chapter 4 verified that there are no TDRS service gaps due to coverage or link dropouts for the user orbits in the hypothetical design study.

Globalstar and Iridium have sufficient capacity to service the scenario data flows in parallel without concern for resource blocking. A prototype NSN commercial service provider modeling and simulation tool is used to compute periods when the free-space communications link cannot be established with the space-based observatories in the hypothetical design study, resulting in data characterizing service gap distributions for each network.[28] A summary of modeling results characterizing the service gaps in the hypothetical transient science scenario for each commercial service provider is provided in Table 19.

Table 19: Summary of results characterizing service gaps for a hypothetical scenario using commercial service providers.

Service Gap Characteristics	Globalstar	Iridium	Units
Service Gap Probability	40%	70%	Percent
Mean Service Gap Period	04:04	06:47	MM:SS
Max Service Gap Period	22:09	26:11	MM:SS
Expected Wait Time Contribution	02:09	04:25	MM:SS

Timeliness Suitability Metric Models

In the hypothetical design study, TDRS network capacity and topology constraints lead to serial servicing of users in the scenario. As a result, the TDRS network timeliness models developed in Chapter 4 differ from those of the commercial networks developed here. The TDRS models developed in Chapter 4 ignore service gaps due to blocking from pre-scheduled user services. The TDRS service gaps due to blocking were characterized in Chapter 3. The TDRS

[28] The space-based observatories in the hypothetical design study are in circular orbits at 400 km altitude and 52 degrees inclination. For a given communications link design, service gap results depend on the relative orbital geometries between user and network service provider nodes. Relative orbital geometries vary over differing timescales and can be affected by natural perturbations and active maneuvers. The goal of the NSN commercial service provider tool is to allow mission and network planners to rapidly explore and identify potential commercial network solutions prior to undertaking higher-fidelity simulations based on proprietary information.

network timeliness models developed in Chapter 4 are extended to incorporate the impacts of TDRS service gaps in this section.

There are two system suitability metrics for timeliness. The first, service start wait time for the i^{th} observatory, $SSWT_i$, measures the queue wait time for each user to begin its servicing by the space data network. Since Globalstar and Iridium space-based nodes have high capacity to service users in parallel and the three co-orbiting user observatories have the same service gap distributions, the SSWT experienced by each user is the same. This simple relationship is expressed mathematically in Eq. 20.

$$SSWT_i = SSWT \tag{20}$$

For the TDRS network, the model for $SSWT_i$ provided in Chapter 4 represents the minimum service start wait time for each observatory in the scenario, $SSWT_{Min}$, since blocking periods due to pre-scheduled user services are ignored in Chapter 4.

The second timeliness metric is the time to complete the scenario service instance, *TTC*. *TTC* measures the expected time interval following initiation of the service access process until the final service execution period ends, completing the data flows in the scenario instance. Since users are serviced in parallel and the command message data volume is small compared to the communications link data transport rate of each commercial service provider, TTC is considered equal to SSWT.[29] This relationship is expressed mathematically in Eq. 21.

$$TTC = SSWT \tag{21}$$

[29] The literature provides observed values for the link data rates of each commercial service provider on the order of 350-1000 bits per second for space-terminated messages for a data transport latency for an 800 bit command message of approximately 1 second [66] [67] [68]. The message data transport latency is included in the inherent latency datasets presented for each commercial service provider.

For the TDRS network, the minimum *TTC* (TTC_{Min}) occurs when there are no blocking periods due pre-scheduled user services and can be calculated directly with the *TTC* model provided in Chapter 4.

For the commercial networks, the minimum *TTC* is equal to the intrinsic latency of each commercial service provider network. This relationship is expressed mathematically in Eq. 22.

$$TTC_{Min} = Intrinsic\_Latency \tag{22}$$

For the commercial networks, the minimum *TTC* occurs during periods when there is end-to-end connectivity between the user's mission control center and the space-based observatory. If a service gap occurs, *TTC* is equal to the intrinsic network latency plus the wait time due to the service gap. The probability of experiencing a service gap due to loss of the free-space communications link for each commercial service provider is provided in Table 19. Since the data flows of the transient science scenario are initiated in response to random events, the expected wait time due to a service gap is equal to half the duration of the mean gap period.[30] The expected wait time due to the distribution of service gap periods for each commercial service provider is provided in Table 19. A model of the expected *TTC* if a service gap occurs (TTC_{Gap}) for the commercial networks is provided in Eq. 23.

$$TTC_{Gap} = Intrinsic\_Latency + Expected\_Gap\_Wait\_Time \tag{23}$$

For the TDRS network, if a service gap due to blocking occurs, the time to complete the scenario (TTC_{Gap}) is equal to the minimum time to complete the scenario (TTC_{Min}) plus the

[30] Roberts et al. and Chapter 3 provide fuller development and discussion of probabilistic wait time contributions for event-driven transient science scenarios.

expected wait time due to the blocking period. The expected wait time of a service gap due to blocking, *Expected_Gap_Wait_Time*, was found to be 2 minutes and 58 seconds in Chapter 3. A model of the expected *TTC* if a service gap occurs (TTC_{Gap}) for the TDRS network is provided in Eq. 24.

$$TTC_{Gap} = TTC_{Min} + Expected\_Gap\_Wait\_Time \quad (24)$$

For the TDRS and commercial networks, the total expected TTC (TTC_{Total}) is a probability weighted sum of the TTC_{Min} and the expected TTC_{Gap}. The probability of a service gap occurrence, P_{Gap}, is provided for each commercial service provider in Table 19 . For the TDRS network P_{Gap} was found to be 16% in Chapter 3.[31] A model for TTC_{Total} is provided in Eq. 25.

$$TTC_{Total} = \left(1 - P_{Gap}\right) * TTC_{Min} + P_{Gap} * TTC_{Gap} \quad (25)$$

Capacity Modeling

There are two system suitability metrics for capacity. First is the number of space data network service instances required per scenario instance, *NSI*. For the simple deterministic data flows of the hypothetical scenario, *NSI* is equal to the total number of space-based observatories in the scenario, *N*. This relationship is expressed mathematically in Eq. 26.

$$NSI = N \quad (26)$$

The second capacity suitability metric is the long-term space data network utilization of a scenario, *NU*. *NU* measures the expected time the network is busy servicing the scenario divided

[31] Per the scenario assumptions defined in Chapter 4, the three user service access requests are treated as a single scheduling trial with a probability of success equal to one minus the blocking probability, P_{Gap}.

by the total network capacity available for all users. The network utilization metric provides a basis for evaluating the impact of the long-term space data network loading that will be imposed on the network service provider for a scenario implementation.

The expected time the space data network is busy servicing the scenario is the product of the transient scientific event mean occurrence rate, *Event_Rate*, and the total expected servicing time to complete the scenario data flows (TTC_{Total}). The total network capacity is computed as the sum of all available servicing channels in the commercial provider network. The literature indicates that the Globalstar and Iridium networks have the capacity to support large numbers of simultaneous user data flows.[32] Each available network channel can support mobile-terminated short message data rates on the order of 350-500 bits per second.

The hypothetical scenario's mean occurrence rate is specified at five events per day. Each transient scientific event deterministically initiates three user data flows, each carrying 800-bit command messages. As a result, the expected long-term network utilization of the scenario is negligible for each commercial service provider.

Throughput Modeling

There are three system suitability metrics for throughput. First, the total data volume transported by the space data network for a scenario instance, *DV*, is the sum of the command data volume sent by each of the command pipelines to their respective observatories, DV_i. The command data volume and number of user command pipelines are deterministic as specified in the scenario implementation and given by Eq. 27 below.

[32] Voss et al. report that each space-based Globalstar network node is designed to support 2,500 data channels. McMahon and Rathburn report that the Iridium SBD services are "optimized for high capacity and efficiency when sending small amounts of data."

$$DV = \sum_{i=1}^{N} DV_i \tag{27}$$

Next, the effective scenario servicing data rate, *EDR*, provides a basis for evaluating the time efficiency of a scenario implementation. *EDR* is calculated as the ratio between the total data volume delivered for a scenario instance, *DV*, and the total expected scenario time to complete, TTC_{Total}. *EDR* is inclusive of queue servicing wait times in a scenario implementation that are not typically captured in traditional network feasibility analyses. *EDR* is calculated by Eq. 28 below.

$$EDR = \frac{DV}{TTC_{Total}} \tag{28}$$

Finally, the long-term scenario data volume throughput, *LTDV*, quantifies the total data volume the space data network is expected to transport while a scenario is in operations. It is computed as the total data volume delivered for a scenario instance, *DV*, and the occurrence rate of the transient scientific event, *Event_Rate*, provided in Eq. 29 below.

$$LTDV = DV * Event\_Rate \tag{29}$$

Financial Cost Modeling

The total lifecycle financial cost of implementing a space data network solution can be estimated based on three component metrics. First, C_{Dev}, is the non-recurring engineering development and commercial procurement costs for the space and ground user-network interfaces. Second, $C_{I\&T}$, is the non-recurring cost for network planning, integration and testing,

culminating in successful data flows across user and network systems. Third, $C_{OM\&S}$, is the recurring cost for operating, maintaining and sustaining data services.

For typical space mission design studies, it is impractical to model all of the factors that influence the network solution lifecycle costs represented by the system suitability cost metrics. Network solution lifecycle costs are also influenced by many external factors and apportioned to users or network service providers based on variety of conditions and arrangements, complicating the cost modeling task [69]. Some aspects of cost modeling are straightforward but lack public data sources. For example, users may procure a space-qualified wireless terminal from several commercial vendors (a contributing factor to C_{Dev}) but pricing data is often marked as proprietary information. Factors which influence the total $C_{OM\&S}$ include costs for operating maintaining and sustaining network assets, which are largely borne by the network service provider, and user operational service costs, C_{Ops}, which are paid by the user to the service provider. In an efficient market, the total revenue a network service provider obtains from user operational service costs (C_{Ops}) balance the total $C_{OM\&S}$ costs borne by the network service provider. Within this context, results from operational service cost models (C_{Ops}) are needed at the "per user" level-of-analysis for user decision and planning purposes and at the "multi-user" scenario level-of-analysis to inform aggregate planning and bulk service procurement decisions by service brokers such as NSN.

The literature provides sufficient data to illustrate at least one model for each of the cost suitability metrics. Operational cost models at the individual user and scenario levels-of-analysis are also developed for each of the three candidate network solutions. All costs are reported in

United States Dollars (USD) in the year identified by the data source. No adjustments are made in the models to normalize for the time varying purchasing power of the dollar.

Iridium Cost Model

Riot et al. provide cost data for the Iridium service plan they selected in March 2020. Activation of a wireless mobile-terminal is reported to be $40. A one-time activation fee to enable Internet Protocol-based ground-to-space (mobile-originated) data flows is also reported but is not strictly required within the scope of the hypothetical design study.33 Activation fees, *Activation_Fees*, are a contributing factor to network integration and test costs.

A model of the network integration and test cost suitability metric, $C_{I\&T}$, is provided by Eq. 30.

$$C_{I\&T} = \sum Activation\_Fees \tag{30}$$

Fees for monthly service plans, *Monthly_Plan_Fee*, are a contributing factor to operational service costs. Riot et al. report the monthly SBD subscription fee for their selected plan to be $19.50. The monthly fee includes usage data volumes up to 136,000 bits (*Monthly_Allowed_Data_Volume*). Monthly data overage fees are billed at $1.40 per 8,000 bits (*Overage_Rate*).

The expected monthly operational service costs for the scenario can be calculated using the long-term data volume throughput of the scenario, *LTDV*, and the corresponding model provided in Eq. 31 and Eq. 32. Note, individual user service costs for participation in the scenario

[33] An additional $500 setup fee to enable Internet Protocol-based space-to-ground data flows is required for Iridium to transport telemetry to mission control centers. Typically, missions require acknowledgement messages verifying receipt of the command onboard the spacecraft. However, the scope of the hypothetical design study is limited to an evaluation of ground-to-space data flows, as they are the limiting factor in timely transient science follow-up observation scenarios.

can also be calculated using only the long-term data volume associated with the individual user within a scenario.

If LTDV < Monthly_Allowed_Data_Volume

$$C_{Ops} = Monthly\_Plan\_Fee \tag{31}$$

If LTDV > Monthly_Allowed_Data_Volume

$$C_{Ops} = Monthly\_Plan\_Fee + Overage\_Rate * (LTDV - Monthly\_Allowed\_Data\_Volume) \tag{32}$$

Additional sources of lifecycle cost data for Iridium were not identified in the literature.

Globalstar Cost Model

Voss et al. report cost data for procurement of Globalstar space-qualified mobile terminals and a 2015 monthly service plan with tiered costs based on data volume usage.[34] However, the cost data reported are for simplex user terminals and space-to-ground (mobile-originated) data services. Ground-to-space (mobile-terminated) data services seem to require procurement of a more complex duplex mobile terminal and likely follow a somewhat different service pricing schedule. Additional sources of lifecycle cost data for Globalstar were not identified in the literature. As a result, costs for the simplex user mobile terminal and operational services are presented here as illustrative models for C_{Dev} and C_{Ops}. Results for these models are not calculated in the hypothetical design study since they do not correspond to mobile-terminated data flows. However, results from these models could be calculated and may be adequate for rough order of magnitude cost estimating and solution feasibility screening purposes until more current and authoritative cost data sources can be identified.

[34] Cost information is not presented in the main conference paper. It appears in the backup section of the associated presentation file [66].

The procurement cost of wireless user terminals, *Terminal_Unit_Cost*, is an important contributor to overall development costs (C_{Dev}). Voss et al. report the procurement cost for an engineering model (non-flight qualified) simplex wireless mobile terminal to be $1,400. The cost of a space-qualified simplex mobile terminal is reported to be $3,600. The space-qualified mobile terminal cost includes some items traditionally accounted for under integration and testing costs in government-led development projects, including electromagnetic interference testing and certification and radio frequency spectrum authorization and licensure costs.

A space-qualified user terminal unit is required for each observatory in the hypothetical design study. In addition, an engineering model unit of the mobile terminal is typically required for each user to support ground and flight system development, testing and sustaining activities. A simple model of the wireless terminal procurement cost for each user in the scenario is provided by Eq. 33.

$$C_{Dev} = \sum Termial\_Unit\_Cost \quad (33)$$

Voss et al. report the 2015 monthly simplex service plan fees according to the tiered pricing structure in Table 20.

Table 20: 2015 Globalstar monthly simplex service plan cost tiers, adapted from Voss et al.

Verified Kilobytes of Data Received	**Cents/Byte, 2015 USD[35] (*Cost_Rate*)**
0 – 360	1.00
361 – 1,800	0.75
1,800 – 3,600	0.50
3,600 – 18,000	0.40
18,000 – 65,318	0.30

The expected monthly operational service costs, C_{Ops}, can be calculated using Table 20 using the appropriate rate, *Cost_Rate*, for the long-term data volume throughput of the scenario, *LTDV*. The units for *LTDV* are in bytes and the costs are in cents in Eq. 34. Note, individual user service costs for participation in the scenario can also be calculated by inputting only the long-term data volume associated with the individual user within a scenario.

$$C_{Ops} = LTDV * Cost\_Rate \tag{34}$$

TDRS Cost Model

The NASA Space Communications and Navigation Mission Operations and Communications Services document provides guidance for developing network lifecycle cost estimates for mission planners including TDRS service rates as of August 2021 [69]. TDRS operational service costs are based on the total number of service minutes provided. Use of the TDRS multiple-access forward service for ground-to-space data flows cost \$15.00 per minute of service. TDRS multiple-access return services cost \$9.00 per minute of service, so a mission seeking to establish a bi-directional link to send and verify receipt of commands onboard the

[35] A 15% discount from the listed rates is offered for academic users.

spacecraft is charged \$24 per minute of service. A 66% discount to the service rates is applied for mission users with a Commercial Space Launch Act agreement.

The scope of the hypothetical design study is limited to an evaluation of ground-to-space data flows, as they are the limiting factor in transient science follow-up observation scenarios. The expected monthly operational service costs for the scenario can be calculated as the product of the transient science event occurrence rate, *Event_Rate*, and the number of service instances per scenario, *NSI*, the per minute TDRS multiple-access forward service cost, *TDRS_MAF_Cost*, and the expected service duration of each data flow service instance in the scenario.

At the maximum TDRS forward data rate of 300,000 bits per second, the theoretical time required to transport the specified command data volume to an observatory in this scenario is less than 0.003 seconds. However, there are implementation specific latencies associated with the observatory's acquisition of the TDRS signal and synchronization of the user flight and ground systems. Additionally, the minimum allowable TDRS service execution period, *Min_SE*, was experimentally verified to be one minute. For the hypothetical design study, it is assumed that the user signal acquisition, software synchronization and data transport processes can be completed within the fixed one-minute TDRS minimum service period. As a result, TDRS service instances have a deterministic execution period of 60 seconds.

The expected monthly operational service costs of the scenario, C_{Ops}, are calculated using the model provided in Eq. 35. Individual user service costs for participation in the scenario are calculated using only the number of network service instances required by the individual user within a scenario.

$$C_{Ops} = Event\_Rate * NSI * Min\_SE * TDRS\_MAF\_Cost \tag{35}$$

Additional sources of lifecycle cost data for TDRS were not identified in the literature.

Step 3: Evaluate system suitability metric outcomes.

In Step 3, the quantitative models developed in Step 2 are executed for the hypothetical design study. Modeling results for the commercial service providers are compared with those of the TDRS network.

The system suitability metrics are intended to provide information to assist both user missions and network service providers in evaluating the feasibility and relative merits of alternative network solutions for a multi-observatory transient science scenario. A summary of results is provided in Table 21.

Table 21 System suitability metric outcomes for network implementation options of a hypothetical transient science scenario.

Suitability Metric Category	Metric	Globalstar	Iridium	TDRS	Units
Timeliness	$SSWT_{Min}$	00:28	00:15	$User_1$ = 06:05 $User_2$ = 07:35 $User_3$ = 09:05	MM:SS
	TTC_{Min}	00:28	00:15	10:05	MM:SS
	TTC_{Total}	01:20	3:10	10:33	MM:SS
Capacity	*NSI*	3	3	3	
	NU	Negligible	Negligible	Negligible	Percent
Throughput	*DV*	2,400	2,400	2,400	Bits per Scenario Instance
	EDR	30	12	3.2	Bits per Second
	LTDV	360,000	360,000	360,000	Bits per Month
Financial Cost	C_{Ops}	No Data[36]	\$59	\$6,750	US Dollars Per Month[37]

For the hypothetical design study, the timeliness outcomes for alternative implementations of ground-to-space data flows are examined. Specifically, the timeliness requirement for the scenario is specified as follows: “The command data shall be delivered as soon as possible to the space-based observatories and must be delivered no later than 10 minutes following initiation of service access processes for space data network services.”

The Globalstar and Iridium commercial service providers satisfy the scenario timeliness requirement by a wide margin, even when accounting for the impact of service gaps (TTC_{Total}).

[36] The data and model for Globalstar cost presented in Step 2 is for space-to-ground (mobile-originated) simplex data flows, whereas the scenario data flows are ground-to-space (mobile-terminated). If pricing for Globalstar ground-to-space data flows is the same, the estimated monthly cost of the scenario data flows is less than one dollar.
[37] Nominal dollar values from the year specified in the data source (i.e., values are not normalized for purchasing power)

The TDRS network is unable to meet the scenario timeliness requirement, even when the impact of service gaps due to blocking are ignored (TTC_{Min}). The dominant wait time contributors to the TDRS TTC_{Min} outcome are the fixed minimum service start lead time (six minutes) and queue servicing delays (three minutes).[38]

The network service providers must evaluate the expected impacts of a new transient science scenario on their ability to satisfy the data transport demands of other users. The expected network utilization metric allows network service providers to compare utilization of a transient science scenario relative to the available network capacity. The Globalstar and Iridium commercial service providers were designed to handle large numbers of randomly initiated user data flows. Although the available capacity of commercial service providers is difficult to determine precisely (and is likely proprietary information) it is unlikely that they have less available capacity than the TDRS network. The TDRS network utilization for the hypothetical design scenario is less than half of one percent of the available capacity. As a result, the network utilization for all network providers is considered negligible.

The effective data rate (*EDR*) provides an indicator of the technical efficiency of a network solution because it normalizes differences in the communications link data rates and expected latencies among different network service providers. Although the TDRS multiple-access forward S-band communications link data rates are up to 1,000 times faster than those of the L-band commercial service providers, this advantage is offset by other factors, as illustrated by the

[38] The TDRS queue servicing delays arise from the as-built TDRS node capacity constraint of one multiple-access forward service instance per node. The TDRS network imposes a fixed inter-service instance delay of 30 seconds. The minimum TDRS service period is 60 seconds. These queue servicing delays can be seen by comparing the service start wait time (SSWT) outcomes for each of the three users in the scenario.

difference in *EDR* outcomes. The commercial service providers are approximately four to 10 times more efficient at transporting the scenario data than TDRS.

The financial cost of monthly data services (C_{Ops}) provides an indicator of the economic efficiency of a network solution. TDRS operational service costs are two orders of magnitude greater than those of the Iridium network for the volume of data transported in this scenario. The TDRS network operational service cost model is based on the number of service minutes provided, whereas the commercial service provider cost models are based on the volume of user data transported.

Discussion of Results

The coordinated observation of transient scientific sources with complimentary ground and space-based observatories requires advances in space mission planning methods and network infrastructure. As the single point of contact for NASA space data network service planning and operations for missions venturing out to two million kilometers from Earth, the NSN seeks to leverage commercial network service provider capabilities to increase the efficiency and robustness of NASA-owned infrastructure.

A hypothetical design study of a transient science follow-up observation scenario was presented in this research to illustrate a new multipoint space data flow design and evaluation method. A set of descriptive SysML architectural models for implementing multipoint space data flows using command pipelines was extended to incorporate commercial network service providers. A set of moderate fidelity quantitative models were developed to predict data flow timeliness and other suitability metrics outcomes for alternative network solutions. An additional

system suitability metric, financial cost, was added to facilitate network solution suitability and screening decisions.

Key Findings

The summary of results in Table 21 indicates that both the Globalstar Short Messaging Service and Iridium Short Burst Data commercial services are viable solutions for disseminating timely commands to space-based observatories, while the as-built TDRS network fails to satisfy the data delivery timeliness requirement in the hypothetical design study.

The minimum time to complete the scenario (TTC_{Min}) timeliness outcomes of both commercial service providers are substantially better than those of the as-built TDRS network which are driven mainly by the TDRS fixed service start lead time. A TDRS scheduling ground rule requires a minimum service start lead time of 10 minutes following a user service request. However, ground testing performed with the TDRS scheduling engineering unit found that service requests were accepted with as little as six minutes of lead time. Nevertheless, the literature demonstrates that the intrinsic latencies of the commercial service providers are regularly less than 30 seconds, and in one instance, as little as 10 seconds.

Scenarios involving many co-orbiting spacecraft also benefit from the large capacity of the commercial service providers. The minimum user service start wait time ($SSWT_{Min}$) metric outcome illustrates the effect of the limited TDRS node servicing capacity, resulting in additional queue wait time due to the fixed TDRS minimum service period (one minute) and inter-user setup lead time (0.5 minute) as users are serviced serially in the hypothetical design study. The free-space command data volume transport time, which depends on the communications link data

rate and light trave time between user and service provider nodes, was negligible for TDRS and the commercial service providers.

TDRS provides 100% orbital coverage to the scenario users, but their randomly initiated command data flows must contend with blocking from pre-scheduled service periods for other TDRS users. By contrast, commercial service providers experience communications signal coverage gaps but have the capacity to handle large numbers of simultaneous users. The TDRS mean wait time due to blocking (2 minutes 58 seconds) was found to be comparable to the mean wait times due to signal coverage gaps predicted by the NSN's prototype commercial service network modeling and simulation tool (2 minutes 9 seconds for Globalstar and 4 minutes 25 seconds for Iridium).[39] The probability weighted average time to complete a scenario instance (TTC_{Total}) is more than three times greater than that of the commercial service providers.

Operational services for the TDRS network are priced based on the number of service minutes used, whereas the commercial service providers use a pricing model based on the volume of user data transported. Since coordinating follow-up observations in a transient science scenario involves an uncertain number of data flows transporting low data volume command messages, a data volume-based pricing model is more advantageous to both users and bulk-buy service brokers such as NSN. This advantage is illustrated in the large difference in expected operational service cost for the scenario data flows ($6,750 for TDRS and $59 for Iridium).

[39] The TDRS blocking probability can be expected to increase as the population of transient science users grows. The magnitude and frequency of Globalstar and Iridium signal coverage gaps depend on user satellite altitude and inclination and other factors. The results presented in this research are based on TDRS blocking data collected in 2019 and for a simulated user orbit of 400 kilometers altitude and 52 degrees inclination.

Demonstration and Practical Considerations

The narrow scope of the hypothetical design study was intended to illustrate the new challenges to space mission design and network planning posed by transient science space systems. An evaluation of TDRS and commercial network solution options for performing one-way command data flows in a transient science scenario was performed. The suitability of the commercial networks for traditional mission critical command and telemetry data flows and for high-volume science data flows was not investigated. The likelihoods and consequences of commercial provider service gaps due to data losses, equipment failures, operator proficiency or other service availability factors were also not considered. Nevertheless, the possibility for improved timeliness outcomes, along with the modest financial costs and spacecraft accommodation requirements, suggest that transient science observatories could fly commercial user terminals for opportunistic transient science data flows as a compliment to a more traditional high-reliability direct-to-Earth or space relay user terminal. In practice, each user mission must consider and balance solutions to satisfy users' traditional network service needs alongside those for one or more transient science observation scenarios. Accordingly, the evaluation method presented in this research may be expanded or applied as an augmentation to existing methods to manage the broader set of traditional mission design and network service planning concerns.

The quantitative models and system suitability metric outcomes presented in this research are intended for design feasibility and network suitability screening without reliance on proprietary information. The models and results rely on publicly available experimental and design data and statistical regressions of computational orbital and communications link

simulation data. Several random variables and probabilistic suitability metric outcomes were characterized and modeled using statistical mean values. This approach has advantages as well as limitations. For example, if the scenario required command data to be delivered within four minutes at least 50% of the time a transient scientific event occurs, system planners could quickly determine that the as-built TDRS network is not viable, based on the experimental evidence and model of TTC_{Min}, and that either commercial service provider could potentially be a viable solution. However, the probability weighted average timeliness model (TTC_{Total}) is not sufficient to determine the likelihood of satisfying the timeliness requirement by the commercial service providers. Characterizing the likelihood of attaining timeliness (and other suitability metric) outcomes could be determined using design-of-experiments empirical techniques or Monte Carlo computational simulation methods. The reliability of the timeliness results would depend on the accuracy of the scenario-dependent service gap model, which could be improved using proprietary communications link simulations or experimental results from Globalstar and Iridium space-based user pathfinder missions.

The latency contributed by the communications network is one part in an end-to-end set of processes that contribute to overall system effectiveness metrics, such as the expected number of events observed during a campaign. Although the TDRS, Globalstar and Iridium networks are suitable for some transient science applications, none of the network service options are able to regularly achieve data delivery timeliness outcomes of less than approximately 30 seconds. To improve the communications network latency further, a TDRS continuous broadcast service could be implemented with upgrades to the TDRS ground systems [36]. Alternatively, a transient science messaging broadcast service could be provided by a

commercial network service provider. High data rate and low latency services could also be achieved by direct user crosslink topologies [70]. Emerging commercial space relay and satellite-based internet broadband service providers may also provide high bandwidth and low latency network solution options.

Conclusion

International researchers in Japan and Australia, as well as an emerging ecosystem of commercial value-added resellers, are demonstrating the potential value of commercial short burst data and messaging service provider networks for space-based user applications. The findings of this research support the proposition that the Globalstar and Iridium commercial service providers can provide complimentary capabilities to the NSN by addressing limitations of the as-built TDRS network to support transient science observation scenarios. Refinements to the NSN prototype commercial service provider quantitative modeling tool are anticipated as more data become publicly available and experience is gained through practical interactions with the transient science mission design community. Such interactions will also contribute to the refinement of transient science user needs and future network service requirements development. The increasing miniaturization, processing efficiency, and software-defined nature of space-qualified wireless user terminals may allow for complimentary multi-network service provider solutions viable for transient science users. Future services available to transient science users may include a TDRS or commercial broadcast service or incorporation of new commercial service providers within the NSN.

Chapter 6

Chapter 6 presents conclusions from the dissertation research.

The study of transient scientific phenomena using time-sensitive measurements from multiple space-based observatories present new challenges to traditional space data network planning methods and current space data network infrastructure systems. This dissertation has defined and completed a series of tasks to address the overarching research question:

How can space data network solutions be designed and evaluated for multi-observatory transient science space systems?

Summary of Findings

In Chapter 3, a systematic investigation of transient science operations involving the Neil Gehrels Swift Observatory, the TDRS space data network and the GCN terrestrial multipoint (publish-subscribe) science messaging application resulted in a descriptive architectural model for transient science space systems codified in the formalisms of SysML structural and behavioral diagrams.

The magnitude and distribution of network service gap periods was determined to be a significant factor for data latency in transient science operations. Two methods for achieving timely network services were defined, experimentally demonstrated and compared. The gap-filling method relies on pre-scheduled service events from complimentary network service providers to reduce the magnitude of the largest network service gaps. The gap-filling method results in guaranteed mean timeliness outcomes at the expense of reserving network service

periods that may go unused. The event-driven method allows on-the-fly access to network services in response to unplanned transient scientific events. The event-driven method results in timeliness outcomes that depend on fixed and random wait times associated with service reservation, service setup and resource blocking by pre-scheduled network users. The network traffic for a transient science observation scenario was determined to be significantly influenced by the occurrence rate of the transient scientific event and attributes of the space data flows required by the scenario, such as the data volume and number of service instances required to execute the scenario. Quantitative models for both methods were developed, and results were calculated for a Swift transient science scenario using the TDRS space data network. The results indicate that the TDRS space data network has adequate capabilities and capacity to allow implementation of both methods by users for transient science applications. Swift has operationalized both methods to perform for rapid follow-up observations of ground-based gravitational wave detections.

In Chapter 4, the transient science space system architectural model was extended to address scenarios involving multiple-space based observatories. The command pipeline and notification pipeline functional reference architectures describe two alternative space-terrestrial network interface design patterns that achieve multipoint (publish-subscribe) data flows. The command pipeline preserves traditional functional partitioning across the space mission user and space data network interface. The notification pipeline requires users to adopt standardized event notifications and allocates greater autonomy to the user mission platform to achieve improved timeliness outcomes and reduced network traffic for transient science scenarios. A quantitative multipoint space data flow modeling method based in queueing theory was defined.

General system suitability metrics for timeliness, throughput and capacity were specified to support the evaluation of alternative network solutions for transient science data flows. A hypothetical design study was performed to apply the command and notification pipeline reference architectures and evaluate alternative TDRS space data flow implementations for a multi-observatory transient science scenario. The merits of a proposed future TDRS broadcast service implementing command or notification pipeline data flows were quantified and discussed.

In Chapter 5, the transient science space system architectural model was extended to incorporate commercial network service providers. Quantitative models for Globalstar and Iridium short messaging data services were developed based on publicly available data sources. Financial cost was added to the set of system suitability metrics. The hypothetical design study in Chapter 4 was extended to compare the relative suitability of the as-built TDRS network with the commercial Globalstar and Iridium network solutions options. The results of the hypothetical design study indicate that both Globalstar and Iridium are viable solutions for disseminating timely commands to space-based observatories. Although the TDRS network has greater coverage of users in low-Earth-orbit, Globalstar and Iridium are capable of achieving more timely outcomes due to the large, fixed minimum service start lead time of the as-built TDRS network. Additionally, both commercial networks have the capacity to handle large numbers of simultaneous user data flows per service provider node, mitigating the impacts of amplified user demand and traffic when a transient scientific event occurs.

Research Contributions

Results from this research address the growing demand for responsive, collaborative, multi-observatory scientific operations concepts by providing a credible model-based systems engineering approach to design multipoint space data flows and evaluate the relative suitability of alternative space communications network solutions. Specifically, the primary contributions of this dissertation are highlighted below.

- An extensive review of the literature and state-of-the-practice for transient science space system planning and operations.
- Identification and quantification of the fixed and probabilistic factors that influence data latency and network traffic for low-Earth-orbit missions using the TDRS, Iridium and Globalstar space data networks.
- A rigorous architecture reference model describing the high-level structure and behavior of transient science space systems, including two functionally distinct design options (command and notification pipelines) available to system planners to implement multipoint space data flows.
- Conceptual development and quantitative models for two network scheduling methods (gap-filling and event-driven) to improve data timeliness outcomes as compared to the state-of-the-practice. The scheduling methods were experimentally demonstrated using the Swift mission and TDRS space data network.
- Conceptual development and quantitative models illustrating the influence of standardized event notification data on system outcomes. Notification data are

common messages, distinct from traditional space mission and engineering datatypes, used to synchronize and orchestrate activities across multiple space-based observatories.

- A novel model-based systems engineering method and set of suitability metrics for designing and evaluating network solutions for transient science space systems.
- Synthesis and demonstration of the model-based systems engineering method and models in a hypothetical transient science space system design study comparing the suitability of as-built and potential future government and commercial space data network solutions.
- Evidence supporting the proposition that the commercial Globalstar and Iridium short messaging data services provide advantages to transient science users due to their lower intrinsic latencies and greater capacity to support simultaneous data flows, as compared to traditional space relay networks such as TDRS.
- Adding to the limited model-based systems engineering application case history for distributed systems.

Results from this research will also be used to inform NASA's technology investment and implementation roadmap for a timely TDRS ground-to-space communications service and the commercialization efforts of the Near Space Network.

A notable practical contribution of this research is that the Swift mission implemented the first known fully autonomous command pipeline in January 2020. As a result, Swift has

increased its expected detection rate of electromagnetic counterparts to ground-observed gravitational waves by greater than 400%.

Future Work

In the future, multipoint data flows transporting command and notification data will be achieved by seamless internetworking between space and terrestrial data networks. The complexity associated with accessing space data network services will be abstracted from the user akin to the experience provided by current terrestrial wireless networks. NASA, in partnership with commercial and international partners, is presently planning such an interoperable network between the Earth and the Moon [71] [72]. Transient science observation scenarios implemented using command and notification pipelines can serve as important pathfinders and drivers to accelerate space-terrestrial internetworking technology infusion and infrastructure development. As space-terrestrial internetworking protocols are adopted by users and network service providers, science messaging applications can be extended directly to space-based observatories, as depicted in below in Fig. 25.

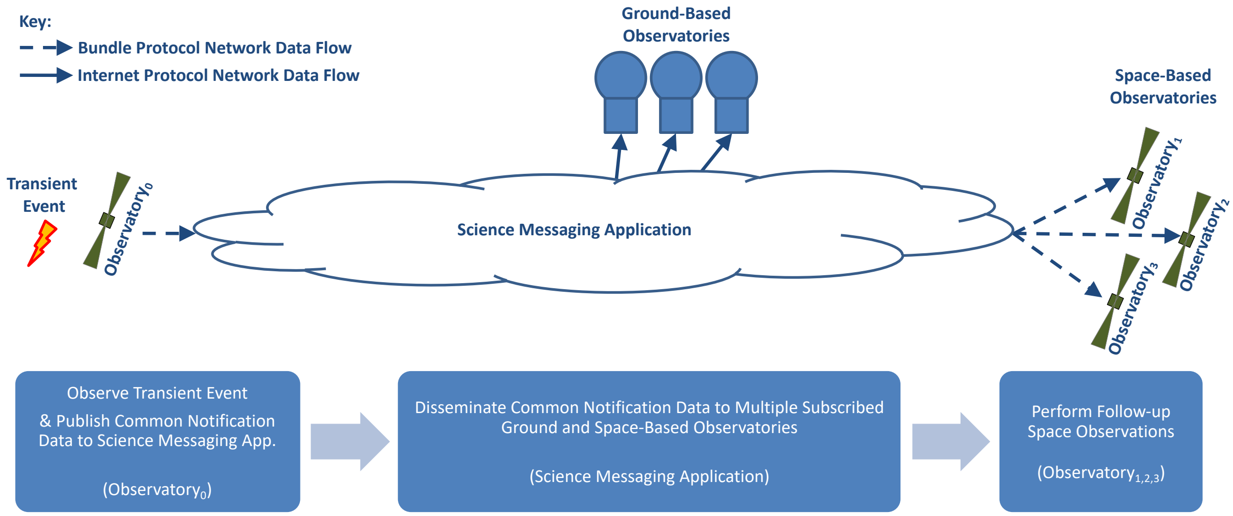


Fig. 25: Operational view of a transient science scenario executed using space-terrestrial internetworking protocols.

Seamless internetworking of space-based observatories with Earth's terrestrial networks could give rise to innovative space operations concepts akin to the "if-this-then-that" automation recipes used in the internet-of-things. Such concepts rely on event-driven machine-to-machine configuration, control, status and synchronization data flows that are not easily captured by traditional space mission design and network planning processes. The command and notification pipeline functional reference architectures, along with the multipoint data flow design and evaluation method presented in this research, can enable a future when space-based observatories become interoperable sensing devices connected by a diverse ecosystem of network service providers.

References

[1] R. M. Sambruna *et al.*, "The NASA Multi-Messenger Astrophysics Science Support Center," *ArXiv210910841 Astro-Ph*, Sep. 2021, Accessed: Sep. 23, 2021. [Online]. Available: http://arxiv.org/abs/2109.10841

[2] P. Mészáros, D. B. Fox, C. Hanna, and K. Murase, "Multi-messenger astrophysics," *Nat. Rev. Phys.*, vol. 1, no. 10, Art. no. 10, Oct. 2019, doi: 10.1038/s42254-019-0101-z.

[3] J. D. Scargle and G. J. Babu, "Ch. 20. Point processes in astronomy: Exciting events in the universe," in *Handbook of Statistics*, vol. 21, Elsevier, 2003, pp. 795–825. doi: 10.1016/S0169-7161(03)21022-1.

[4] N. Gehrels *et al.*, "The Swift Gamma-Ray Burst Mission," *Astrophys. J.*, vol. 611, no. 2, p. 1005, Aug. 2004, doi: 10.1086/422091.

[5] B. P. Abbott *et al.*, "Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A," *Astrophys. J.*, vol. 848, no. 2, p. L13, Oct. 2017, doi: 10.3847/2041-8213/aa920c.

[6] D. Selva, A. Golkar, O. Korobova, I. L. i Cruz, P. Collopy, and O. L. de Weck, "Distributed Earth Satellite Systems: What Is Needed to Move Forward?," *J. Aerosp. Inf. Syst.*, vol. 14, no. 8, pp. 412–438, Aug. 2017, doi: 10.2514/1.I010497.

[7] S. A. Chien *et al.*, "Automated Volcano Monitoring Using Multiple Space and Ground Sensors," *J. Aerosp. Inf. Syst.*, vol. 17, no. 4, pp. 214–228, Jan. 2020, doi: 10.2514/1.I010798.

[8] M. Patterson, "Streaming Data from the Universe with Apache Kafka," *Confluent*. https://www.confluent.io/blog/streaming-data-from-the-universe-with-apache-kafka (accessed Feb. 18, 2021).

[9] D. J. Israel, G. W. Heckler, and R. J. Menrad, "Space Mobile Network: A Near Earth Communications and Navigation Architecture," Big Sky, MT, United States, Jan. 2016, pp. 1–7. doi: doi: 10.1109/AERO.2016.7500669.

[10] J. Garvin, "Future Integrated Robotic and Exploration Science," presented at the Space Communications and Navigation Users, Plans, and Requirements Working Group, NASA Headquarters, Washington, DC, Aug. 20, 2015.

[11] S. Kaplan and B. Guarino, "Scientists detect gravitational waves from a new kind of nova, sparking a new era in astronomy," *Washington Post*, Oct. 16, 2017. Accessed: Aug. 16, 2019. [Online]. Available: https://www.washingtonpost.com/news/speaking-of-

science/wp/2017/10/16/scientists-detect-gravitational-waves-from-a-new-kind-of-nova-sparking-a-new-era-in-astronomy/

[12] S. Kaplan, “In a cosmic first, scientists detect ‘ghost particles’ from a distant galaxy,” *Washington Post*, Jul. 12, 2018. Accessed: Aug. 16, 2019. [Online]. Available: https://www.washingtonpost.com/news/speaking-of-science/wp/2018/07/12/in-a-cosmic-first-scientists-detect-ghostly-neutrinos-from-a-distant-galaxy/

[13] S. Potter, “NASA Missions Catch First Light from a Gravitational wave Event,” *NASA*, Oct. 16, 2017. http://www.nasa.gov/press-release/nasa-missions-catch-first-light-from-a-gravitational-wave-event (accessed Jul. 24, 2019).

[14] M. W. E. Smith *et al.*, “The Astrophysical Multimessenger Observatory Network (AMON),” *Astropart. Phys.*, vol. 45, pp. 56–70, May 2013, doi: 10.1016/j.astropartphys.2013.03.003.

[15] N. Gehrels, J. K. Cannizzo, J. Kanner, M. M. Kasliwal, S. Nissanke, and L. P. Singer, “Galaxy Strategy for LIGO-Virgo Gravitational Wave Counterpart Searches,” *Astrophys. J.*, vol. 820, no. 2, p. 136, Mar. 2016, doi: 10.3847/0004-637X/820/2/136.

[16] “The Swift Technical Handbook Version 12.0.” https://swift.gsfc.nasa.gov/proposals/tech_appd/swiftta_v12/swiftta_v12.html (accessed Jul. 24, 2019).

[17] “About the Gamma Coordinates Network (GCN)/Transient Astronomy Network (TAN).” https://gcn.gsfc.nasa.gov/about.html (accessed Jul. 24, 2019).

[18] G. Hohpe and B. Woolf, *Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions*, 1st ed. Addison-Wesley Professional. Part of the Addison-Wesley Signature Series (Fowler) series., 2003. Accessed: May 08, 2019. [Online]. Available: http://www.informit.com/store/enterprise-integration-patterns-designing-building-9780321200686?ranMID=24808

[19] “Near Space Network,” *Exploration and Space Communications Division, Goddard Space Flight Centerb*. https://esc.gsfc.nasa.gov/projects/NSN (accessed Jan. 18, 2022).

[20] J. R. Wertz, D. F. Everett, and J. J. Puschell, Eds., *Space Mission Engineering: The New SMAD*, vol. 28. Hawthorne, CA: Microcosm Press, 2011.

[21] W. Braun and T. McKenzie, “CLASS: A Comprehensive Satellite Link Simulation Package,” *IEEE J. Sel. Areas Commun.*, vol. 2, no. 1, pp. 129–137, Jan. 1984, doi: 10.1109/JSAC.1984.1146033.

[22] A. Kossiakoff, W. N. Sweet, S. Seymour, and S. M. Biemer, *Systems Engineering Principles and Practice*, 2 edition. Wiley-Interscience, 2011.

[23] *NASA Systems Engineering Handbook*, SP-2016-6105 Rev. 2., vol. 1 & 2. Office of the Chief Engineer, NASA Headquarters, Washington, D.C., 2016. [Online]. Available: https://nodis3.gsfc.nasa.gov/displayDir.cfm?t=NPR&c=7123&s=1B

[24] J. M. Borky and T. H. Bradley, *Effective Model-Based Systems Engineering*. Springer International Publishing, 2019. Accessed: Mar. 18, 2019. [Online]. Available: https://www.springer.com/gp/book/9783319956688

[25] S. Friedenthal, A. Moore, and R. Steiner, *A Practical Guide to SysML: The Systems Modeling Language*, 2nd ed. Elsevier, 2012. doi: 10.1016/C2010-0-66331-0.

[26] J. A. Estefan, "Survey of Model-Based Systems Engineering (MBSE) Methodologies." International Council on Systems Engineering (INCOSE), 2008.

[27] NASA Office of the Chief Engineer, "NASA Space Mission Architecture Framework (SMAF) Handbook for Uncrewed Space Missions." Mar. 11, 2021. [Online]. Available: file:///Users/cjrober5/Desktop/2021-03-11_nasa-hdbk-1005_baseline_0.pdf

[28] J. Sokol, "Minisatellite surge spurs downlink infrastructure," *Science*, vol. 357, no. 6358, pp. 1342–1343, Sep. 2017, doi: 10.1126/science.357.6358.1342.

[29] M. E. Brown, M. L. Carroll, and V. M. Escobar, "User needs and assessing the impact of low latency NASA Earth observation data availability on societal benefit," *Space Policy*, vol. 30, no. 3, Part A, pp. 135–137, Aug. 2014, doi: 10.1016/j.spacepol.2014.05.002.

[30] M. Sanchez Net, "Support of latency-sensitive space exploration applications in future space communication systems," Thesis, Massachusetts Institute of Technology, 2017. Accessed: Oct. 30, 2018. [Online]. Available: http://dspace.mit.edu/handle/1721.1/112458

[31] "Swift Fact Sheet." Accessed: Jan. 15, 2020. [Online]. Available: https://swift.gsfc.nasa.gov/about_swift/Sci_Fact_Sheet.pdf

[32] T. A. Gitlin and W. Horne, "The NASA Space Network Demand Access System (DAS)," presented at the SpaceOps 2002 Conference, Houston, Texas, Oct. 2002. doi: 10.2514/6.2002-T2-50.

[33] T. M. Hackett and M. D. Johnston, "Investigating a Demand Access Scheduling Paradigm for NASA's Deep Space Network," in *11th International Workshop on Planning and Scheduling for Space (IWPSS)*, Berkeley, CA, Jul. 2019, pp. 51–60.

[34] T. M. Hackett, M. Johnston, and S. G. Bilen, "Spacecraft Block Scheduling for NASA's Deep Space Network," in *2018 SpaceOps Conference*, American Institute of Aeronautics and Astronautics. doi: 10.2514/6.2018-2578.

[35] D. J. Israel, F. Davis, and J. Marquart, "A DTN-Based Multiple Access Fast Forward Service for the NASA Space Network," in *2011 IEEE Fourth International Conference on Space*

Mission Challenges for Information Technology, Aug. 2011, pp. 61–65. doi: 10.1109/SMC-IT.2011.19.

[36] G. W. Heckler, C. Gramling, J. Valdez, and P. Baldwin, “TDRSS Augmentation Service for Satellites,” in *SpaceOps 2016 Conference*, 0 vols., American Institute of Aeronautics and Astronautics, 2016. doi: 10.2514/6.2016-2467.

[37] Human Exploration and Operations Mission Directorate, “NASA Policy Directive (NPD) 8074.1: Management and Utilization of NASA’s Space Communication and Navigation Infrastructure,” Aug. 11, 2009. https://nodis3.gsfc.nasa.gov/displayDir.cfm?t=NPD&c=8074&s=1 (accessed Apr. 11, 2020).

[38] C. Guidorzi, S. Dichiara, F. Frontera, R. Margutti, A. Baldeschi, and L. Amati, “A Common Stochastic Process Rules Gamma-Ray Burst Prompt Emission and X-Ray Flares,” *Astrophys. J.*, vol. 801, no. 1, p. 57, Mar. 2015, doi: 10.1088/0004-637X/801/1/57.

[39] H. Tijms, *A First Course in Stochastic Models*, 1st ed. John Wiley & Sons, Ltd, 2003. doi: 10.1002/047001363X.

[40] “Laser Interferometer Gravitational wave Observatory (LIGO) Frequently Asked Questions,” *LIGO Lab | Caltech*. https://www.ligo.caltech.edu/page/faq (accessed May 24, 2020).

[41] S. D. Barthelmy *et al.*, “GRB Coordinates Network (GCN): A status report,” *AIP Conf. Proc.*, vol. 526, no. 1, pp. 731–735, Sep. 2000, doi: 10.1063/1.1361631.

[42] E. J. Birrane, D. J. Copeland, and M. G. Ryschkewitsch, “The path to space-terrestrial internetworking,” in *2017 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, Oct. 2017, pp. 134–139. doi: 10.1109/WiSEE.2017.8124906.

[43] D. Israel, B. Edwards, J. Hayes, W. Knopf, A. Robles, and L. Braatz, “The Benefits of Delay/Disruption Tolerant Networking (DTN) for Future NASA Science Missions,” presented at the 70th International Astronautical Congress (IAC), Washington, D.C., Oct. 2019. doi: IAC-19, B2,7,3,x53530.

[44] C. Araguz, E. Bou-Balust, and E. Alarcón, “Applying autonomy to distributed satellite systems: Trends, challenges, and future prospects,” *Syst. Eng.*, vol. 21, no. 5, pp. 401–416, 2018, doi: https://doi.org/10.1002/sys.21428.

[45] E. Crawley, B. Cameron, and D. Selva, *System Architecture: Strategy and Product Development for Complex Systems*, 1st edition. Boston: Pearson, 2015.

[46] C. J. Roberts *et al.*, “Evaluation of Timely Communications Access Methods Using NASA Space Network,” *AIAA J. Aerosp. Inf. Syst.*, vol. 18, no. 6, p. 14, Jun. 2021, doi: https://doi.org/10.2514/1.I010897.

[47] "Asynchronous Message Service," *Consult. Comm. Space Data Syst. CCSDS*, vol. CCSDS 735.1-B-1, Blue Book, no. September, 2011, p. 142, 2011.

[48] The Consultative Committee for Space Data Systems (CCSDS), "Security Architecture for Space Data Systems," CCSDS 351.0-M-1, Nov. 2012. [Online]. Available: https://public.ccsds.org/Pubs/351x0m1.pdf

[49] M. Sanchez Net, I. del Portillo, B. Cameron, and E. F. Crawley, "Architecting Information Security Services for Federated Satellite Systems," *J. Aerosp. Inf. Syst.*, vol. 14, no. 8, pp. 439–450, Aug. 2017, doi: 10.2514/1.I010425.

[50] "Recommendations on the Selection of End-to-End Space Internetworking Protocol." Interagency Operations Advisory Group (IOAG) Space Internetworking Strategy Group (SISG).

[51] B. S. Blanchard and W. J. Fabrycky, *Systems Engineering and Analysis*, 5 edition. Boston: Pearson, 2010.

[52] K. Lynaugh *et al.*, "S-Band Transponder Multi-Network Compatibility, Space Environment and Radiation Testing," *Small Satell. Conf.*, Aug. 2020, [Online]. Available: https://digitalcommons.usu.edu/smallsat/2020/all2020/176

[53] A. Tohuvavohu, J. A. Kennea, J. DeLaunay, D. M. Palmer, S. B. Cenko, and S. Barthelmy, "Gamma-Ray Urgent Archiver for Novel Opportunities (GUANO): Swift/BAT Event Data Dumps on Demand to Enable Sensitive Subthreshold GRB Searches," *Astrophys. J.*, vol. 900, no. 1, p. 35, Aug. 2020, doi: 10.3847/1538-4357/aba94f.

[54] E. Birrane, P. Chimento, D. Copeland, and B. Haberman, "An Architecture for Advanced Networking Technology for Integrating Communications in Space," in *2020 IEEE Aerospace Conference*, Mar. 2020, pp. 1–7. doi: 10.1109/AERO47225.2020.9172813.

[55] National Academies of Sciences, Engineering, and Medicine, *Pathways to Discovery in Astronomy and Astrophysics for the 2020s*. Washington, DC: The National Academies Press, 2021. doi: 10.17226/26141.

[56] D. Baird, "NASA to Commercialize Near-Earth Communications Services," *NASA.gov*, Oct. 26, 2020. http://www.nasa.gov/feature/Goddard/2020/nasa-to-commercialize-near-earth-communications-services (accessed Dec. 03, 2021).

[57] B. Younes, "SCaN the Future: Commercializing Near-Earth Communications," NASA HQ, Virtual Presentation Event, Oct. 29, 2020. Accessed: Dec. 03, 2021. [Online]. Available: https://www.nasa.gov/sites/default/files/atoms/files/scan_the_future_-_integrated_master_briefing_export.pdf

[58] G. Karpati, J. Martin, M. Steiner, and K. Reinhardt, "The integrated mission design center (imdc) at nasa goddard space flight center," in *2003 IEEE Aerospace Conference*

Proceedings (Cat. No.03TH8652), Mar. 2003, vol. 8, p. 8_3657-8_3667. doi: 10.1109/AERO.2003.1235549.

[59] T. Gitlin and K. Walyus, “NASA’s Space Network Ground Segment Sustainment Project Preparing for the Future,” p. 14.

[60] Y. Hata *et al.*, “Flight Model Development of the AGU Remote Innovative CubeSat Alert System - ARICA,” *Small Satell. Conf.*, Aug. 2021, [Online]. Available: https://digitalcommons.usu.edu/smallsat/2021/all2021/41

[61] D. Messier, “Epsilon Launches 9 Japanese Technology Demonstration Satellites – Parabolic Arc.” http://parabolicarc.com/2021/11/11/epsilon-launches-9-japanese-technology-demonstration-satellites/ (accessed Nov. 16, 2021).

[62] D. Lambeth, “SpIRIT of Australia Set To Launch By 2022,” Jun. 23, 2020. https://spaceaustralia.com/news/spirit-australia-set-launch-2022 (accessed Nov. 29, 2021).

[63] K. Khan, “Data Communication With A Nano-satellite Using Satellite Personal Communication Networks (s-pcns),” University of Central Florida, Department of Electrical Engineering and Computer Science, 2008. [Online]. Available: https://stars.library.ucf.edu/etd/3588

[64] C. Rodriguez, H. Boiardt, and S. Bolooki, “CubeSat to commercial intersatellite communications: Past, present and future,” in *2016 IEEE Aerospace Conference*, Mar. 2016, pp. 1–15. doi: 10.1109/AERO.2016.7500525.

[65] R. Mearns and M. Trenti, “Near real-time telecommand solutions for CubeSats: State of the art and applications to the SkyHopper mission,” presented at the 17th Australian Space Research Conference, University of Sydney, Aug. 2018. Accessed: Aug. 19, 2019. [Online]. Available: http://arxiv.org/abs/1808.06746

[66] H. Voss, J. Dailey, M. Orvis, A. White, and S. Brandle, “Globalstar Link: From Reentry Altitude and Beyond,” Aug. 2016.

[67] V. J. Riot, L. M. Simms, and D. Carter, “Lessons Learned Using Iridium to Communicate with a CubeSat in Low Earth Orbit,” *J. Small Satell.*, vol. Vol. 10, no. No. 1, pp. 995–1006, Feb. 2021.

[68] M. M. McMahon and R. Rathburn, “Measuring Latency in Iridium Satellite Constellation Data Services,” Naval Academy Annapolis MD Department of Computer Science, Jun. 2005. Accessed: Oct. 28, 2021. [Online]. Available: https://apps.dtic.mil/sti/citations/ADA464192

[69] Space Communications and Navigation Program, “Mission Operations and Communications Services.” NASA Headquarters, Washington, D.C., Aug. 13, 2021. [Online].

Available: https://explorers.larc.nasa.gov/2021APMIDEX/pdf_files/SCaN-MOCS-0001-Rev%20_4__Final.pdf

[70] J. E. Velazco, "Omnidirectional Optical Communicator," in *2019 IEEE Aerospace Conference*, Mar. 2019, pp. 1–6. doi: 10.1109/AERO.2019.8741924.

[71] D. J. Israel *et al.*, "LunaNet: a Flexible and Extensible Lunar Exploration Communications and Navigation Infrastructure," in *2020 IEEE Aerospace Conference*, Mar. 2020, pp. 1–14. doi: 10.1109/AERO47225.2020.9172509.

[72] "Draft LunaNet Interoperability Specification," National Aeronautics and Space Administration, LN-IS Baseline V001, Sep. 2021. Accessed: Sep. 30, 2021. [Online]. Available: https://esc.gsfc.nasa.gov/static-files/Draft%20LunaNet%20Interoperability%20Specification%20Final.pdf